

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:16:49 ; Search time 7.30435 Seconds
(without alignments)
530.147 Million cell updates/sec

Title: US-09-856-319B-4_COPY_34_264

Perfect score: 1252

Sequence: 1 IVNGENAVPGSWPQVSLQD.....AMVTRVSKFTWVQVMAYN 231

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 106657 seqs, 16763532 residues

Total number of hits satisfying chosen parameters: 106657

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	756	60.4	192	10	US-09-925-297-529
2	755.5	60.3	263	10	US-09-888-615-96
3	464.5	37.1	251	9	US-09-961-721-4
4	464.5	37.1	791	9	US-09-967-386-1
5	464.5	37.1	810	10	US-09-946-893-2
6	462	36.9	812	9	US-09-333-323-1
7	462	36.9	812	10	US-09-788-142-1
8	462	36.9	812	10	US-09-761-120-1
9	462	36.9	812	10	US-09-873-676-81
10	459.5	36.7	270	10	US-09-923-779-152
11	457.5	36.5	249	9	US-09-961-721-5
12	457.5	36.5	273	10	US-09-925-297-529
13	455.5	36.4	453	9	US-09-978-295A-69
14	455.5	36.4	453	9	US-09-978-697-69
15	455.5	36.4	453	9	US-09-978-192A-69
16	455.5	36.4	453	12	US-10-052-586-64
17	455	36.3	248	10	US-09-925-301-1017
18	455	36.3	327	10	US-09-804-156-16
19	455	36.3	327	10	US-09-946-633-8

20	455	36.3	454	10	US-09-888-615-103	Sequence 103, App
21	448.5	35.8	343	10	US-09-948-094-2	Sequence 2, Appli
22	438.5	35.0	226	9	US-09-961-721-6	Sequence 6, Appli
23	438.5	35.0	226	9	US-10-045-367A-5	Sequence 5, Appli
24	438.5	35.0	226	10	US-09-910-151-6	Sequence 6, Appli
25	430.5	34.4	393	9	US-10-012-896-934	Sequence 934, App
26	430.5	34.4	393	10	US-09-759-143-934	Sequence 934, App
27	430.5	34.4	393	10	US-09-780-669-934	Sequence 934, App
28	430.5	34.4	393	10	US-09-822-827-934	Sequence 934, App
29	430.5	34.4	492	9	US-10-012-896-932	Sequence 932, App
30	430.5	34.4	492	10	US-09-759-143-932	Sequence 932, App
31	430.5	34.4	492	10	US-09-780-669-932	Sequence 932, App
32	430.5	34.4	492	10	US-09-822-827-932	Sequence 932, App
33	428.5	34.2	492	9	US-10-012-896-895	Sequence 895, App
34	428.5	34.2	492	10	US-09-759-143-895	Sequence 895, App
35	428.5	34.2	492	10	US-09-780-669-895	Sequence 895, App
36	428.5	34.2	492	10	US-09-822-827-895	Sequence 895, App
37	428.5	34.2	492	10	US-09-879-752-14	Sequence 14, Appl
38	427.5	34.1	384	9	US-09-981-353-23	Sequence 23, Appl
39	424	33.9	320	10	US-09-888-615-90	Sequence 90, Appl
40	423.5	33.8	457	10	US-09-888-615-110	Sequence 110, App
41	423	33.8	223	10	US-09-910-071-14	Sequence 14, Appl
42	422.5	33.7	414	10	US-09-820-893-69	Sequence 69, Appl
43	422.5	33.7	480	10	US-09-820-893-108	Sequence 108, App
44	422	33.7	1169	9	US-09-870-759-126	Sequence 126, App
45	421.5	33.7	283	10	US-09-988-975A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-925-297-529

; Sequence 529, Application US/09925297

; Patent No. US20020081659A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA105

; CURRENT APPLICATION NUMBER: US/09/925,297

; CURRENT FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05989

; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 928

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 529

; LENGTH: 192

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-925-297-529

Query Match

Best Local Similarity 60.4%; Score 756; DB 10; Length 192;

Matches 135; Conservative 16; Mismatches 9; Indels 0; Gaps 0;

QY 1 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVTAACQVTPGRHFVVLGEYDRS 60

||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 30 IVNGENAVJGSPWQVSLQDSSGFHFCGSLISQSWVTAACQVTPGRHFVVLGEYDRS 89

||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 61 SNAEPQVLSIARAIHPNNANTMNDLTLLKLSAPRYTAQVSPVCLASTNEALPSGL 120

||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 90 SNAEPLOVLSVRAIHPNSWNTMNDVTLKLSAPQYTRISPVCLASSNEALTEGL 149

||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 121 TCVTGWRISGVNGVTPARLQOVLPVTVNOCROYGA 160

||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 150 TCVTGWRISGVNGVTPARLQOVLPVTVNOCROYGS 189

||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 2

US-09-888-615-96

; Sequence 96, Application US/09888615

```

; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: FLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR FILING DATE: 2001-06-26
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 96
; LENGTH: 263
; TYPE: PRP
; ORGANISM: Homo sapiens
US-09-888-615-96

Query Match      60.3%; Score 755.5; DB 10; Length 263;
Best Local Similarity 58.4%; Pred. No. 7.3e-62;
Matches 135; Conservative 35; Mismatches 60; Indels 1; Gaps 1;

Qy 1 IVNGENAVPGSWPQVSLQDNTGFHFCGGLISPNWVYTAHCQVTPGRHFVVLGEYDRS 60
Db 34 IVNGEDAVPGSWPQVSLQDNTGFHFCGGLISPDWVYTAHCQVTPGRHFVVLGEYDRS 92
Qy 61 SNAEPQVLSIARAITHPNMNTMNDLTLKLLASPARYTAQVSPVCLASTNEALPSGL 120
Db 93 SDENIQVLIKAKVFNKPSILTVNNDITLLKLTAPARETSQVSAVCLPSADDDFPAGT 152
Qy 121 TCVTTGGRISGVGNVTPARLQOVVPLVTVNQCROYNGARITDAMICAGSGAGSCQGD 180
Db 153 LCATTGKTKYNKATPDKLQQAALPLLSNAECKSKMGRITDVTMICAGASGVSSCMGD 212
Qy 181 SGGPLVCQKGNWTLVIGSVGKNCNIQAPAMTRYSKFSTWNOVMAYN 231
Db 213 SGGPLVCQKGNWTLVIGSVGKNCNITTPAVYARVTKLIPWQKILAN 263

RESULT 3
US-09-961-721-4
; Sequence 4, Application US/09961721
; Patent No. US20020156005A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: m32404, A NOVEL HUMAN TRYPSIN AND USES
; FILE REFERENCE: 10448-096001
; CURRENT APPLICATION NUMBER: US/09/961,721
; CURRENT FILING DATE: 2001-09-24
; PRIOR FILING DATE: 2000-09-24
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 251
; TYPE: PRP
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus sequence
US-09-961-721-4

Query Match      37.1%; Score 464.5; DB 9; Length 251;
Best Local Similarity 44.2%; Pred. No. 2.5e-35;
Matches 111; Conservative 26; Mismatches 79; Indels 35; Gaps 12;

Qy 9 PGSW--PKQVSLQDNTG----FHFCGGLISPNWVYTAHC-----QVTPGRHF-----VV 53
Db 1 PGFSGPQVSLQVRSRGGGRKHFHFCGGLISENVWLTAAHCVCASAPASSVRSLSVR 60

```

```

Qy 54 LGEYDRSSNAEPVQVLSIARA-IHPNWNANTM-----NNDLTLKLLASPARYTAQ-VSP 106
Db 61 LGEHNLSTECTEQKFDVKKTIIVHPNYPNPTLDNGAYDNDIALKLLKSPGVTLGDIVRP 120
Qy 107 VCLASTNEALPSGLTCTVTTGGRISGVGNVTPARLQOVVPLVTVNQCROYNGARITDAMICAG-- 161
Db 121 ICLPSASDLPGVGTCTVSGWGRRTKNLGLSDTLQEVVVPVSVRETCSRSAEYEGGTDK 180
Qy 162 ---ITDAMICAGG-SGASSCGDSSGGLVCOQGN---TWVLIGIVSGTKNC-NIQAPAM 213
Db 181 VEFVTDNMICAGALGGKDACQDGGGLVCSGDRGDRWELGVISWGSYGCARGNKG 240
Qy 214 YTRVSKFSTWI 224
Db 241 YTRVSSYLDWI 251

RESULT 4
US-09-967-386-1
; Sequence 1, Application US/09967386
; Patent No. US20020159992A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Henkin, Jack
; APPLICANT: Davidson, Donald J.
; TITLE OF INVENTION: ANTANGIOGENIC POLYPEPTIDES AND METHODS
; FILE REFERENCE: 6738-US-02
; CURRENT APPLICATION NUMBER: US/09/967,386
; CURRENT FILING DATE: 2001-09-26
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 791
; TYPE: PRP
; ORGANISM: Homo sapiens
US-09-967-386-1

Query Match      37.1%; Score 464.5; DB 9; Length 791;
Best Local Similarity 41.8%; Pred. No. 1e-34;
Matches 100; Conservative 29; Mismatches 93; Indels 17; Gaps 7;

Qy 1 IVNGENAVPGSWPQVSLQDNTGFHFCGGLISPNWVYTAHC---QVTPGRHFVVLGEY 57
Db 562 VVGCVAHPHSWPQVSLRTRFGMHFCGGLISPEWVLTAAHLEKSPRSPSYKVLGAH 621
Qy 58 DRSSNAEP-VQVLSIARAITHPNMNTMNDLTLKLLASPARYTAQVSPVCLASTNEAL 116
Db 622 -QEVNLEPHVQIEVSRLEPT-----RKDIALKLLSPAVITDKVIPACLPSPNVV 674
Qy 117 PSGLTCTVTTGGRISGVGNVTPARLQOVVPLVTVNQCROYNGARITDAMICAG--GS 172
Db 675 ADRTCEFTIGWGTQGTFGA--GLLKEAQLPVIENKVCNRYEFLNGRVQSTELCAGHLA 732
Qy 173 GASSCGDSSGGLVCOQGNWTLVIGIVSGTKNCNIQAPAMTRYSKFSTWNOVMAYN 231
Db 733 GTDSCQDSSGGLVCOQGNWTLVIGIVSGTKNCNIQAPAMTRYSKFSTWNOVMAYN 791

RESULT 5
US-09-946-893-2
; Sequence 2, Application US/09946893
; Patent No. US20020072494A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yihai
; TITLE OF INVENTION: Materials and methods relating to endothelial cell growth
; FILE REFERENCE: Mawburn
; CURRENT APPLICATION NUMBER: US/09/946,893
; CURRENT FILING DATE: 2001-09-05

```



```

;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: Plasminogen
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-788-142-1

Query Match          36.9%; Score 462; DB 10; Length 812;
Best Local Similarity 41.0%; Pred. No. 1.8e-34;
Matches 98; Conservative 32; Mismatches 93; Indels 16; Gaps 6;

QY 1 IVNGENAVPGSWPQVSLQDN-TGFHFCCGSLSPNWWVYTAHC---QVTPGRHFVVLGE 56
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 582 VVGCVANPHSWPQISLRTFTGQHFCCGTLTAPWVLTAAHCLKSSRPFYKVLGA 641
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 57 YDRSSNAEPQVLSIARITHPNWNTMNDLTLLKLSAPARYTAQVSPVCLASTNEAL 116
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 642 HEEYIRGLDQVLSIAKLILEPN-----NRDIALKLSRPATITDKVIPACLPSPNYW 695
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 117 PSGLTCVTTGWGRISGVNTPARLQOVLPVTVNQCR--QYWGARITDAMICAG--GS 172
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 696 ADRTICYITGWGETQGTGCA--GRLKEAQLPVIENKVCNRVEYLNRRVKSTELCAGQLAG 753
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 173 GASSCOGDSGGPLVCOKGNTWVLIGIVSGTKNCNTQAPAMYTRVSKFSTWVNOVMAYN 231
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 754 GVDSCOGDSGGPLVCPEKDKYILQGVTSWGLGCARPKNKPGVYRVSRFVDWIERMRNN 812
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 8
US-09-761-120-1
; Sequence 1, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasminogen
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09/761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 812
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
; NAME/KEY: misc.feature
; OTHER INFORMATION: Plasminogen
US-09-761-120-1

Query Match          36.9%; Score 462; DB 10; Length 812;
Best Local Similarity 41.0%; Pred. No. 1.8e-34;
Matches 98; Conservative 32; Mismatches 93; Indels 16; Gaps 6;

QY 1 IVNGENAVPGSWPQVSLQDN-TGFHFCCGSLSPNWWVYTAHC---QVTPGRHFVVLGE 56
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 582 VVGCVANPHSWPQISLRTFTGQHFCCGTLTAPWVLTAAHCLKSSRPFYKVLGA 641
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 57 YDRSSNAEPQVLSIARITHPNWNTMNDLTLLKLSAPARYTAQVSPVCLASTNEAL 116
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 642 HEEYIRGLDQVLSIAKLILEPN-----NRDIALKLSRPATITDKVIPACLPSPNYW 695
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 117 PSGLTCVTTGWGRISGVNTPARLQOVLPVTVNQCR--QYWGARITDAMICAG--GS 172
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 696 ADRTICYITGWGETQGTGCA--GRLKEAQLPVIENKVCNRVEYLNRRVKSTELCAGQLAG 753
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 173 GASSCOGDSGGPLVCOKGNTWVLIGIVSGTKNCNTQAPAMYTRVSKFSTWVNOVMAYN 231
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 754 GVDSCOGDSGGPLVCPEKDKYILQGVTSWGLGCARPKNKPGVYRVSRFVDWIERMRNN 812
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
US-09-873-676-81
; Sequence 81, Application US/09873676
; Patent No. US20020077289A1
; GENERAL INFORMATION:
; APPLICANT: Macdonald, Nicholas J.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Angiotensin and Endostatin Binding Proteins and Met
; FILE REFERENCE: 05213-0378 (43170-259333)
; CURRENT APPLICATION NUMBER: US/09/873,676
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 60/209,065
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: US 60/289,387
; PRIOR FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 81
; LENGTH: 812
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-873-676-81

Query Match          36.9%; Score 462; DB 10; Length 812;
Best Local Similarity 41.0%; Pred. No. 1.8e-34;
Matches 98; Conservative 32; Mismatches 93; Indels 16; Gaps 6;

QY 1 IVNGENAVPGSWPQVSLQDN-TGFHFCCGSLSPNWWVYTAHC---QVTPGRHFVVLGE 56
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 582 VVGCVANPHSWPQISLRTFTGQHFCCGTLTAPWVLTAAHCLKSSRPFYKVLGA 641
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 57 YDRSSNAEPQVLSIARITHPNWNTMNDLTLLKLSAPARYTAQVSPVCLASTNEAL 116
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 642 HEEYIRGLDQVLSIAKLILEPN-----NRDIALKLSRPATITDKVIPACLPSPNYW 695
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 117 PSGLTCVTTGWGRISGVNTPARLQOVLPVTVNQCR--QYWGARITDAMICAG--GS 172
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 696 ADRTICYITGWGETQGTGCA--GRLKEAQLPVIENKVCNRVEYLNRRVKSTELCAGQLAG 753
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 173 GASSCOGDSGGPLVCOKGNTWVLIGIVSGTKNCNTQAPAMYTRVSKFSTWVNOVMAYN 231
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 754 GVDSCOGDSGGPLVCPEKDKYILQGVTSWGLGCARPKNKPGVYRVSRFVDWIERMRNN 812
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
US-09-923-779-152
; Sequence 152, Application US/09923779
; Patent No. US2002007621A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.553
; CURRENT APPLICATION NUMBER: US/09/923,779
; CURRENT FILING DATE: 2001-08-06
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 152
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-923-779-152

Query Match          36.7%; Score 459.5; DB 10; Length 270;

```

Query Match 36.7%; Score 459.5; DB 10; Length 270;


```
Best Local Similarity 42.3%; Pred. No. 7.8e-35;
Matches 102; Conservative 38; Mismatches 88; Indels 13; Gaps 9;

QY 1 IVNGENAVPGSWPQVLSIQ-DNTG--FHFCGGLSPNNVVTAAHCQVTPGRHFVVLGEY 57
Db 29 VVHGEDAVPYSWPQVLSIQEKSGSYHTCGGSLIAPDWWVTAGHCISRDLYQVVLGEY 88
QY 58 DRSSNAEPVQVLSI--ARAITHPNNTM--NNDLTLLKLASPARYTAQVSPVCLASTN 113
Db 89 NLAVKEGPEQVPIINSEELFVHPLNRSVCACGNDIALIKLSRQAGDVLASLPAG 148
QY 114 EALPSGLTCVTGWRISGNGVNTPARLQVVLPLVTNQCQ--YWGARITDAMICAGG 171
Db 149 DILPNKTCYITGWRGLYNGPL-PDKLQARLPVVDYKHCRRNWNWGSTVKKTWVCAGG 207
QY 172 SGASSCGDSGGPLVC-QKGNVTWVLIGVSWGK-NCN-IQAPMYTRVSKFSTWVNM 228
Db 208 YIRSGCNGDSGGPLNCPTEGQVHGVTFSVSGFCNFIWKPVTFTVRSFIDWIEETI 267
QY 229 A 229
Db 268 A 268

RESULT 11
US-09-961-721-5
; Sequence 5, Application US/09961721
; Patent No. US20020156005A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: m32404, A NOVEL HUMAN TRYPSIN AND USES
; FILE REFERENCE: 10448-096001
; CURRENT APPLICATION NUMBER: US/09/961,721
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/235,023
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus sequence
US-09-961-721-5

Query Match 36.5%; Score 457.5; DB 9; Length 249;
Best Local Similarity 44.1%; Pred. No. 1.1e-34;
Matches 108; Conservative 25; Mismatches 79; Indels 33; Gaps 11;

QY 13 PQVSLQDNTG----FHFCGGLSPNNVVTAAHC-----QVTPGRHF-----VVLGEYDR 59
Db 5 PQVSLQVRSRGGGRKHFHFCGGLSPNNVVTAAHC-----VVLGEYDR 64
QY 60 SSNAEPVQVLSI--ARAITHPNNTM--NNDLTLLKLASPARYTAQ-VSPVCLAST 112
Db 65 SLTEGTEQKDFVKKTIIVHPNTNPOTLDNGAYDNDIALLKSPGVTGLDTRVRLCLPSA 124
QY 113 NEALPSGLTCVTGWRISGNGVNTPARLQVVLPLVTNQCQ--YWGARITDAMICAGG 164
Db 125 SSDLPVGTCTVSGWRGRTKGLSLDYLQEVVVPVSRSETCRSAYEGGTDKVEFVTD 184
QY 165 AMICAGG-SGASSCGDSGGPLVCQKGN---TWVLIGVSWGKNC-NIQAPMYTRVSK 219
Db 185 NMICAGGKDKACQDSGGPLVCSDGNDRGRWELVGIWGSYGCGARGNKPGVTVRVSS 244
QY 220 FSTWI 224
Db 245 YLDWI 249

RESULT 12
US-09-925-297-695
; Sequence 695, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 695
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (27)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (28)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-695

Query Match 36.5%; Score 457.5; DB 10; Length 273;
Best Local Similarity 42.3%; Pred. No. 1.2e-34;
Matches 102; Conservative 37; Mismatches 89; Indels 13; Gaps 9;

QY 1 IVNGENAVPGSWPQVLSIQ-DNTG--FHFCGGLSPNNVVTAAHCQVTPGRHFVVLGEY 57
Db 32 VVHGEDAVPYSWPQVLSIQEKSGSYHTCGGSLIAPDWWVTAGHCISRDLYQVVLGEY 91
QY 58 DRSSNAEPVQVLSI--ARAITHPNNTM--NNDLTLLKLASPARYTAQVSPVCLASTN 113
Db 92 NLAVKEGPEQVPIINSEELFVHPLNRSVCACGNDIALIKLSRQAGDVLASLPAG 151
QY 114 EALPSGLTCVTGWRISGNGVNTPARLQVVLPLVTNQCQ--YWGARITDAMICAGG 171
Db 152 DILPNKTCYITGWRGLYNGPL-PDKLQARLPVVDYKHCRRNWNWGSTVKKTWVCAGG 210
QY 172 SGASSCGDSGGPLVC-QKGNVTWVLIGVSWGK-NCN-IQAPMYTRVSKFSTWVNM 228
Db 211 YIRSGCNGDSGGPLNCPTEGQVHGVTFSVSGFCNFIWKPVTFTVRSFIDWIEETI 270
QY 229 A 229
Db 271 A 271

RESULT 13
US-09-978-295A-69
; Sequence 69, Application US/09978295A
; Patent No. US20020156006A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
```

; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Hillan, Kenneth J.
 ; APPLICANT: Kljavin, Ivar J.
 ; APPLICANT: Kuo, Sophia S.
 ; APPLICANT: Napier, Mary A.
 ; APPLICANT: Pan, James;
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Shelton, David L.
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE OF INVENTION: Acids Encoding the Same
 ; FILE REFERENCE: P2630Plc11
 ; CURRENT APPLICATION NUMBER: US/09/978,295A
 ; CURRENT FILING DATE: 2001-10-15
 ; PRIOR APPLICATION NUMBER: 09/918585
 ; PRIOR FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: 60/062250
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/064249
 ; PRIOR FILING DATE: 1997-11-03
 ; PRIOR APPLICATION NUMBER: 60/065311
 ; PRIOR FILING DATE: 1997-11-13
 ; PRIOR APPLICATION NUMBER: 60/066364
 ; PRIOR FILING DATE: 1997-11-21
 ; PRIOR APPLICATION NUMBER: 60/077450
 ; PRIOR FILING DATE: 1998-03-10
 ; PRIOR APPLICATION NUMBER: 60/077632
 ; PRIOR FILING DATE: 1998-03-11
 ; PRIOR APPLICATION NUMBER: 60/077641
 ; PRIOR FILING DATE: 1998-03-11
 ; PRIOR APPLICATION NUMBER: 60/077649
 ; PRIOR FILING DATE: 1998-03-11
 ; PRIOR APPLICATION NUMBER: 60/077791
 ; PRIOR FILING DATE: 1998-03-12
 ; PRIOR APPLICATION NUMBER: 60/078004
 ; PRIOR FILING DATE: 1998-03-13
 ; PRIOR APPLICATION NUMBER: 60/078886
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/078936
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/078910
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/078939
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/079294
 ; PRIOR FILING DATE: 1998-03-25
 ; PRIOR APPLICATION NUMBER: 60/079656
 ; PRIOR FILING DATE: 1998-03-26
 ; PRIOR APPLICATION NUMBER: 60/079664
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079689
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079663
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079728
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079786
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079920
 ; PRIOR FILING DATE: 1998-03-30
 ; PRIOR APPLICATION NUMBER: 60/079923
 ; PRIOR FILING DATE: 1998-03-30
 ; PRIOR APPLICATION NUMBER: 60/080105
 ; PRIOR FILING DATE: 1998-03-31
 ; PRIOR APPLICATION NUMBER: 60/080107
 ; PRIOR FILING DATE: 1998-03-31
 ; PRIOR APPLICATION NUMBER: 60/080165
 ; PRIOR FILING DATE: 1998-03-31
 ; PRIOR APPLICATION NUMBER: 60/080194
 ; PRIOR FILING DATE: 1998-03-31
 ; PRIOR APPLICATION NUMBER: 60/080327
 ; PRIOR FILING DATE: 1998-04-01
 ; PRIOR APPLICATION NUMBER: 60/080328
 ; PRIOR FILING DATE: 1998-04-01
 ; PRIOR APPLICATION NUMBER: 60/080333
 ; PRIOR FILING DATE: 1998-04-01
 ; PRIOR APPLICATION NUMBER: 60/080334
 ; PRIOR FILING DATE: 1998-04-01
 ; PRIOR APPLICATION NUMBER: 60/081070
 ; PRIOR FILING DATE: 1998-04-08
 ; PRIOR APPLICATION NUMBER: 60/081049
 ; PRIOR FILING DATE: 1998-04-08
 ; PRIOR APPLICATION NUMBER: 60/081071
 ; PRIOR FILING DATE: 1998-04-08
 ; PRIOR APPLICATION NUMBER: 60/081195
 ; PRIOR FILING DATE: 1998-04-08
 ; PRIOR APPLICATION NUMBER: 60/081203
 ; PRIOR FILING DATE: 1998-04-09
 ; PRIOR APPLICATION NUMBER: 60/081229
 ; PRIOR FILING DATE: 1998-04-09
 ; PRIOR APPLICATION NUMBER: 60/081955
 ; PRIOR FILING DATE: 1998-04-15
 ; PRIOR APPLICATION NUMBER: 60/081817
 ; PRIOR FILING DATE: 1998-04-15
 ; PRIOR APPLICATION NUMBER: 60/081819
 ; PRIOR FILING DATE: 1998-04-15
 ; PRIOR APPLICATION NUMBER: 60/081952
 ; PRIOR FILING DATE: 1998-04-15
 ; PRIOR APPLICATION NUMBER: 60/081838
 ; PRIOR FILING DATE: 1998-04-15
 ; PRIOR APPLICATION NUMBER: 60/082568
 ; PRIOR FILING DATE: 1998-04-21
 ; PRIOR APPLICATION NUMBER: 60/082569
 ; PRIOR FILING DATE: 1998-04-21
 ; PRIOR APPLICATION NUMBER: 60/082704
 ; PRIOR FILING DATE: 1998-04-22
 ; PRIOR APPLICATION NUMBER: 60/082796
 ; PRIOR FILING DATE: 1998-04-23
 ; PRIOR APPLICATION NUMBER: 60/083336
 ; PRIOR FILING DATE: 1998-04-27
 ; PRIOR APPLICATION NUMBER: 60/083322
 ; PRIOR FILING DATE: 1998-04-28
 ; PRIOR APPLICATION NUMBER: 60/083392
 ; PRIOR FILING DATE: 1998-04-29
 ; PRIOR APPLICATION NUMBER: 60/083495
 ; PRIOR FILING DATE: 1998-04-29
 ; PRIOR APPLICATION NUMBER: 60/083496
 ; PRIOR FILING DATE: 1998-04-29
 ; PRIOR APPLICATION NUMBER: 60/083499
 ; PRIOR FILING DATE: 1998-04-29
 ; PRIOR APPLICATION NUMBER: 60/083545
 ; PRIOR FILING DATE: 1998-04-29
 ; PRIOR APPLICATION NUMBER: 60/083554
 ; PRIOR FILING DATE: 1998-04-29
 ; PRIOR APPLICATION NUMBER: 60/083558
 ; PRIOR FILING DATE: 1998-04-29
 ; PRIOR APPLICATION NUMBER: 60/083559
 ; PRIOR FILING DATE: 1998-04-29
 ; PRIOR APPLICATION NUMBER: 60/083500
 ; PRIOR FILING DATE: 1998-04-29
 ; PRIOR APPLICATION NUMBER: 60/083742
 ; PRIOR FILING DATE: 1998-04-30
 ; PRIOR APPLICATION NUMBER: 60/084366
 ; PRIOR FILING DATE: 1998-05-05

;
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500

;
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 36.4%; Score 455.5; DB 9; Length 453;

Best Local Similarity 37.9%; Pred. No. 3.4e-34;

Matches 89; Conservative 39; Mismatches 96; Indels 11; Gaps 6;

QY 1 IVNGENAVPGSNPWQVSLQDNTGFHCGGSLISPNNVYTAHC---QVTPGRHFVVLGEY 57
Db 217 IVGGNMSLLSQWPQASLQFQ-GYHLCGGSVITPLIITAAHCYDYLPLPKSWTIOVGLV 275
QY 58 DRSSNAEPQVLSIARATHPNKNANTMNDLTLKLASPARYTAQVSPVCLASTNEALP 117
Db 276 SLDDNPAPSHL--VEKIVYHSKYKPKRLGNDIALMKLAGPLTFNEMIQPVCLPENSENFP 333
QY 118 SGLTCVTGTGWRISGVGNVTPARLQOVPLVTVNOC--ROYMGARITDAMICAG--GSG 173
Db 334 DGKVCWTSWGATEDGDASPV-LNHAAYPLISNKNCHNEDYVGGIISPSMLCAGYLTGG 392
QY 174 ASSCQDGGGGLVCQKGNWTWVLIGIVSWGFKNCNIQAPAMYTRVSKFSTWINQVM 228
Db 393 VDSQCGDSGGLVCQERRLWKLVGATSFIGICAEVKNKPGVYTRVTSFLDWIHEQM 447

RESULT 15

US-09-978-192A-69

; Sequence 69, Application US/09978192A

; Patent No. US20020177553A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Baker Kevin P.

APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC9
CURRENT APPLICATION NUMBER: US/09/978,192A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29

[illegible]

•

result No.	Query				DB		ID	Description
	Score	Match	Length					
1	621.5	50.9	231	2	US-09-027-337-6	Sequence 6, Appli		
2	621.5	50.9	231	4	US-09-644-600-6	Sequence 6, Appli		
3	616.5	50.5	230	4	US-08-944-483-62	Sequence 62, Appl		
4	609	49.9	229	2	US-08-557-146-13	Sequence 13, Appl		
5	609	49.9	229	2	US-09-154-344-13	Sequence 13, Appl		
6	607.5	49.8	228	1	US-08-278-091-10	Sequence 10, Appl		
7	607.5	49.8	228	1	US-08-483-859-10	Sequence 10, Appl		
8	607.5	49.8	228	1	US-08-472-173-10	Sequence 10, Appl		
9	607.5	49.8	228	2	US-08-487-167-10	Sequence 10, Appl		
10	607.5	49.8	228	2	US-08-482-816-10	Sequence 10, Appl		
11	607.5	49.8	228	2	US-08-296-149-10	Sequence 10, Appl		
12	607.5	49.8	228	2	US-08-801-499-10	Sequence 10, Appl		
13	607.5	49.8	228	2	US-08-615-271-10	Sequence 10, Appl		
14	607.5	49.8	228	3	US-09-074-660-10	Sequence 10, Appl		
15	607.5	49.8	228	3	US-09-074-659-10	Sequence 10, Appl		
16	607.5	49.8	228	3	US-09-106-468-10	Sequence 10, Appl		
17	607.5	49.8	228	4	US-09-106-466A-10	Sequence 10, Appl		
18	607.5	49.8	228	4	US-09-106-467-10	Sequence 10, Appl		
19	424	34.7	319	4	US-09-386-649-12	Sequence 12, Appl		
20	419	34.3	241	4	US-08-944-483-59	Sequence 59, Appl		
21	416.5	34.1	242	4	US-08-944-483-57	Sequence 57, Appl		
22	408.5	33.5	454	4	US-09-518-046-2	Sequence 2, Appli		
23	402	32.9	814	1	US-08-750-711-1	Sequence 1, Appli		
24	399	32.7	841	4	US-08-944-483-60	Sequence 60, Appl		
25	398	32.6	791	1	US-08-643-219-1	Sequence 1, Appli		
26	398	32.6	791	3	US-08-851-350-1	Sequence 1, Appli		
27	397.5	32.6	299	4	US-08-944-483-66	Sequence 66, Appl		

RESULT 2
US-09-644-600-6
; Sequence 6, Application US/09644600
; Patent No. 6451500
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotooshi

;; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease
;; FILE REFERENCE: Overexpressed in Carcinomas
;; CURRENT APPLICATION NUMBER: US/09/644,600

;; PRIOR FILING DATE: 2000-08-23
;; PRIOR APPLICATION NUMBER: 09/421,213

;; PRIOR FILING DATE: 1999-10-20
;; PRIOR APPLICATION NUMBER: 09/027,337

;; PRIOR FILING DATE: 1998-02-20
;; NUMBER OF SEQ ID NOS: 98

;; SEQ ID NO 6

;; LENGTH: 231

;; TYPE: PRT

;; ORGANISM: Homo sapiens

;; FEATURE:

;; OTHER INFORMATION: Chymotrypsin

US-09-644-600-6

Query Match 50.9%; Score 621.5; DB 4; Length 231;

Best Local Similarity 56.8%; Pred. No. 5e-53;

Matches 113; Conservative 31; Mismatches 54; Indels 1; Gaps 1;

QY 33 RYNGENAVLGSWPQVSLQDSSGFHFCGSLISQSWVYTAHCNVSPGRHFVVLGEYDR 92

DB 1 RYNGEDAVPGSWPQVSLQDKTGFHFCGSLISEDWVYVTAHCGVRTS-DVVVAGEFDQ 59

QY 93 SNAEPLQVLSVSRATPHSPNSMTTNNNDVTLKLASPAQYTRISPVCLASSNEALTEG 152

DB 60 GSDEENIQVLKAKYKFNKPSILTNNNDITLLKATPARFQTSVAVCLPSADDDFFAG 119

QY 153 LTCVTTGWRGLSGVGNVTPAHQQVALPLVTNQCQRYWDSITDSMICAGGAGSSCOG 212

DB 120 TLCATTGWGKTKYANKTPDKLQQAALPLLSNAECKKSGRRITDVMICAGASGVSSCMG 179

QY 213 DSGGPLVCQKGNWTWLVGI 231

DB 180 DSGGPLVCQKGNWTWLVGI 198

RESULT 3

US-08-944-483-62

Sequence 62, Application US/08944483

Patent No. 6232456

GENERAL INFORMATION:

APPLICANT: COHEN, MAURICE

APPLICANT: COLPITTS, TRACEY L.

APPLICANT: FRIEDMAN, PAULA N.

APPLICANT: GRANADOS, EDWARD N.

APPLICANT: KLASS, MICHAEL R.

APPLICANT: RUSSELL, JOHN C.

APPLICANT: STEWART, KENT D.

APPLICANT: STROUPE, STEVEN D.

TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS

TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES

TITLE OF INVENTION: OF THE PROSTATE

NUMBER OF SEQUENCES: 76

CORRESPONDENCE ADDRESS:

STREET: 100 Abbott Park Road

CITY: Abbott Park

STATE: IL

COUNTRY: USA

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/944,483

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:

NAME: Becker, Cheryl L.

REGISTRATION NUMBER: 35,441

REFERENCE/DOCKET NUMBER: 6183.US.01

TELECOMMUNICATION INFORMATION:

TELEPHONE: 847/935-1729

TELEFAX: 847/938-2623

TELEX:

INFORMATION FOR SEQ ID NO: 62:

SEQUENCE CHARACTERISTICS:

LENGTH: 230 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: No. 6232456e

US-08-944-483-62

Query Match 50.5%; Score 616.5; DB 4; Length 230;

Best Local Similarity 56.6%; Pred. No. 1.5e-52;

Matches 112; Conservative 31; Mismatches 54; Indels 1; Gaps 1;

QY 34 IVNGENAVLGSWPQVSLQDSSGFHFCGSLISQSWVYTAHCNVSPGRHFVVLGEYDRS 93

DB 1 IVNGEDAVPGSWPQVSLQDKTGFHFCGSLISEDWVYVTAHCGVRTS-DVVVAGEFDQ 59

QY 94 SNAEPLQVLSVSRATPHSPNSMTTNNNDVTLKLASPAQYTRISPVCLASSNEALTEGL 153

DB 60 GSDEENIQVLKAKYKFNKPSILTNNNDITLLKATPARFQTSVAVCLPSADDDPPAGT 119

QY 154 LTCVTTGWRGLSGVGNVTPAHQQVALPLVTNQCQRYWDSITDSMICAGGAGSSCOGD 213

DB 120 TLCATTGWGKTKYANKTPDKLQQAALPLLSNAECKKSGRRITDVMICAGASGVSSCMGD 179

QY 214 DSGGPLVCQKGNWTWLVGI 231

DB 180 DSGGPLVCQKGNWTWLVGI 197

RESULT 4

US-08-557-146-13

Sequence 13, Application US/08557146

Patent No. 5834290

GENERAL INFORMATION:

APPLICANT: Egelrud, Torbjorn

APPLICANT: Hansson, Lennart

TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic

TITLE OF INVENTION: Enzyme (SCCE)

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036-2787

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/557,146

FILING DATE: 14-DEC-1995

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Sterner, Richard J.

REGISTRATION NUMBER: 35,372

REFERENCE/DOCKET NUMBER: 1103326-181

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 819-8783

TELEFAX: (212) 354-8113

INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 229 amino acids
 TYPE: amino acids
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: polypeptide
 US-08-557-146-13

Query Match 49.9%; Score 609; DB 2; Length 229;
 Best Local Similarity 57.1%; Pred. No. 8.2e-52;
 Matches 113; Conservative 30; Mismatches 53; Indels 2; Gaps 2;
 QY 34 IVNGENAVLGSWPQVSLQDSSGFHFCGSLISQSWVYTAACHNCVSPGRHFWVLGEYDRS 93
 DB 1 IVNGEDAVPGSWPQVSLQDRTGFHFCGSLISEDWVYTAACHGVRTS-DVVVAGEFFDQG 59
 QY 94 SNAEPQLVLSVSRATHPSWNTMNDVTLLKLSPAQYTRISPVCLASSNEALTEGL 153
 DB 60 SDEENIQVLKIAKVFKNPKFSILTVNNDITLLKLATPAREFQTVSAVCLPSADDDFPAGT 119
 QY 154 TCVITGWRGLSGVGNVTPAHLQVVALPLVTYNQCRQYWDSSITDSMICAGGAGASSCGD 213
 DB 120 LCATTGWGKTKYNANKTPDKLQQAALPLLSNAECKKSGRRITDYMICA-GAGVSSCMGD 178
 QY 214 SGGPLVCOCKGNTWVLGI 231
 DB 179 SGGPLVCOCKGAWTLVGI 196

RESULT 5
 US-09-154-344-13
 Sequence 13, Application US/09154344
 Patent No. 5981256
 GENERAL INFORMATION:
 APPLICANT: Egelrud, Torbjorn
 APPLICANT: Hansson, Lennart
 TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
 TITLE OF INVENTION: Enzyme (SCCE)
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: White & Case, Patent Department
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036-2787

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/154,344
 FILING DATE: 16-SEP-1998
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/557,146
 FILING DATE: 14-DEC-1995
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Snetter, Richard J.
 REGISTRATION NUMBER: 35,372
 REFERENCE/DOCKET NUMBER: 1103326-181
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 819-8783
 TELEFAX: (212) 354-8113
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 229 amino acids
 TYPE: amino acids
 STRANDEDNESS: single
 TOPOLOGY: linear

MOLECULE TYPE: polypeptide
 US-09-154-344-13

Query Match 49.9%; Score 609; DB 2; Length 229;
 Best Local Similarity 57.1%; Pred. No. 8.2e-52;
 Matches 113; Conservative 30; Mismatches 53; Indels 2; Gaps 2;
 QY 34 IVNGENAVLGSWPQVSLQDSSGFHFCGSLISQSWVYTAACHNCVSPGRHFWVLGEYDRS 93
 DB 1 IVNGEDAVPGSWPQVSLQDRTGFHFCGSLISEDWVYTAACHGVRTS-DVVVAGEFFDQG 59
 QY 94 SNAEPQLVLSVSRATHPSWNTMNDVTLLKLSPAQYTRISPVCLASSNEALTEGL 153
 DB 60 SDEENIQVLKIAKVFKNPKFSILTVNNDITLLKLATPAREFQTVSAVCLPSADDDFPAGT 119
 QY 154 TCVITGWRGLSGVGNVTPAHLQVVALPLVTYNQCRQYWDSSITDSMICAGGAGASSCGD 213
 DB 120 LCATTGWGKTKYNANKTPDKLQQAALPLLSNAECKKSGRRITDYMICA-GAGVSSCMGD 178
 QY 214 SGGPLVCOCKGNTWVLGI 231
 DB 179 SGGPLVCOCKGAWTLVGI 196

RESULT 6
 US-08-278-091-10
 Sequence 10, Application US/08278091
 Patent No. 5506139
 GENERAL INFORMATION:
 APPLICANT: LOOSMORE, Sheena M
 APPLICANT: YANG, Yan-Ping
 APPLICANT: CHONG, Pele
 APPLICANT: OOMEN, Raymond P.
 APPLICANT: KLEIN, Michel H.
 TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
 TITLE OF INVENTION: Reduced Protease Activity
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sim & McBurney
 STREET: Suite 701, 330 University Avenue
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada
 ZIP: M5G 1R7

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/278,091
 FILING DATE: 21-JUL-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Stewart, Michael I
 REGISTRATION NUMBER: 24,973
 REFERENCE/DOCKET NUMBER: 1038-371
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (416) 595-1155
 TELEFAX: (416) 595-1163
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 228 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

US-08-278-091-10
 Sequence 10, Application US/08278091
 Patent No. 5506139
 GENERAL INFORMATION:
 APPLICANT: LOOSMORE, Sheena M
 APPLICANT: YANG, Yan-Ping
 APPLICANT: CHONG, Pele
 APPLICANT: OOMEN, Raymond P.
 APPLICANT: KLEIN, Michel H.
 TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
 TITLE OF INVENTION: Reduced Protease Activity
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sim & McBurney
 STREET: Suite 701, 330 University Avenue
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada
 ZIP: M5G 1R7

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/278,091
 FILING DATE: 21-JUL-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Stewart, Michael I
 REGISTRATION NUMBER: 24,973
 REFERENCE/DOCKET NUMBER: 1038-371
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (416) 595-1155
 TELEFAX: (416) 595-1163
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 228 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

Query Match 49.8%; Score 607.5; DB 1; Length 228;
 Best Local Similarity 55.6%; Pred. No. 1.1e-51;
 Matches 110; Conservative 36; Mismatches 49; Indels 3; Gaps 3;
 QY 34 IVNGENAVLGSWPQVSLQDSSGFHFCGSLISQSWVYTAACHNCVSPGRHFWVLGEYDRS 93

Db 1 IVNGEAVPGSWPQVSLQDKTGFHFCGSLINENWVVTAAHCGVTTSDVYVAGEFDQG 59
 QY 94 SNAEPLOVLSVRAITHPSWNTTMMNDVTLKSLASPAQYTTTRISPVCLASNEALTEGL 153
 Db 60 SSKSEKIOKIAKVFKNKYNLSLTINDITLLKSLTAASFSQTVSAVCLPSASDDFAAGT 119
 QY 154 TCVTTGWRGLSGVGNVTPAHLOQVALPLVTVNOCRQYWDSSITDSMICAGGAGASCQGD 213
 Db 120 TCVTTGWRGLSGVGNVTPAHLOQVALPLVTVNOCRQYWDSSITDSMICAGGAGASCQGD 213
 QY 214 SGGPLVCQKGNWTWLVGI 231
 Db 178 SGGPLVCCKNGAWTLVGI 195

RESULT 7

US-08-483-859-10
 ; Sequence 10, Application US/08483859
 ; Patent No. 5656436
 ; GENERAL INFORMATION:
 ; APPLICANT: LOOSMORE, Sheena M.
 ; APPLICANT: YANG, Yan-Ping
 ; APPLICANT: CHONG, Pele
 ; APPLICANT: COMEN, Raymond P.
 ; APPLICANT: KLEIN, Michel H.
 ; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
 ; TITLE OF INVENTION: Reduced Protease Activity
 ; NUMBER OF SEQUENCES: 23
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Sim & McBurney
 ; STREET: Suite 701, 330 University Avenue
 ; CITY: Toronto
 ; STATE: Ontario
 ; COUNTRY: Canada
 ; ZIP: M5G 1R7

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/483,859
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/296,149
 ; FILING DATE: 26-AUG-1994

REGISTRATION NUMBER: 24,973
 ; REFERENCE/DOCKET NUMBER: 1038-495 MIS:vg
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (416) 595-1155
 ; TELEFAX: (416) 595-1155
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 228 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-483-859-10

Query Match 49.8%; Score 607.5; DB 1; Length 228;
 Best Local Similarity 55.6%; Pred. No. 1.1e-51;
 Matches 110; Conservative 36; Mismatches 49; Indels 3; Gaps 3;
 QY 34 IVNGENAVLGSWPQVSLQDKTGFHFCGSLINENWVVTAAHCGVTTSDVYVAGEFDQG 93
 Db 1 IVNGEAVPGSWPQVSLQDKTGFHFCGSLINENWVVTAAHCGVTTSDVYVAGEFDQG 59

QY 94 SNAEPLOVLSVRAITHPSWNTTMMNDVTLKSLASPAQYTTTRISPVCLASNEALTEGL 153
 Db 60 SSKSEKIOKIAKVFKNKYNLSLTINDITLLKSLTAASFSQTVSAVCLPSASDDFAAGT 119
 QY 154 TCVTTGWRGLSGVGNVTPAHLOQVALPLVTVNOCRQYWDSSITDSMICAGGAGASCQGD 213
 Db 120 TCVTTGWRGLSGVGNVTPAHLOQVALPLVTVNOCRQYWDSSITDSMICAGGAGASCQGD 213
 QY 214 SGGPLVCQKGNWTWLVGI 231
 Db 178 SGGPLVCCKNGAWTLVGI 195

RESULT 8

US-08-472-173-10
 ; Sequence 10, Application US/08472173
 ; Patent No. 5665353
 ; GENERAL INFORMATION:
 ; APPLICANT: LOOSMORE, Sheena M.
 ; APPLICANT: YANG, Yan-Ping
 ; APPLICANT: CHONG, Pele
 ; APPLICANT: COMEN, Raymond P.
 ; APPLICANT: KLEIN, Michel H.
 ; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
 ; TITLE OF INVENTION: Reduced Protease Activity
 ; NUMBER OF SEQUENCES: 23
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Sim & McBurney
 ; STREET: Suite 701, 330 University Avenue
 ; CITY: Toronto
 ; STATE: Ontario
 ; COUNTRY: Canada
 ; ZIP: M5G 1R7

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/472,173
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/296,149
 ; FILING DATE: 26-AUG-1994

REGISTRATION NUMBER: 24,973
 ; REFERENCE/DOCKET NUMBER: 1038-493 MIS:vg
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (416) 595-1155
 ; TELEFAX: (416) 595-1155
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 228 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-472-173-10

Query Match 49.8%; Score 607.5; DB 1; Length 228;
 Best Local Similarity 55.6%; Pred. No. 1.1e-51;
 Matches 110; Conservative 36; Mismatches 49; Indels 3; Gaps 3;
 QY 34 IVNGENAVLGSWPQVSLQDKTGFHFCGSLINENWVVTAAHCGVTTSDVYVAGEFDQG 93
 Db 1 IVNGEAVPGSWPQVSLQDKTGFHFCGSLINENWVVTAAHCGVTTSDVYVAGEFDQG 59
 QY 94 SNAEPLOVLSVRAITHPSWNTTMMNDVTLKSLASPAQYTTTRISPVCLASNEALTEGL 153

Db	60	SSSEKIQKUKIAKFKNSKYNLSITINDITLLKLSAAFSQTVSAVCLPSASDQFAAGT	119
QY	154	TCVTTGWRGLSGVGNVTPAHLQVALVTVNOCROYWDSSTTDSMICAGGAGASSCGD	213
Db	120	TCVTTGCG-LTRYAN-TPDRLQASLPLLSNTNCKYNGTKIDAMICAGASGVSSCGMD	177
QY	214	SGGPLVCQKGNWTWVLIGI	231
Db	178	SGGPLVCKKNGAMTWLVGI	195

RESULT 9
US-08-487-167-10
; Sequence 10, Application US/08487167
; Patent No. 5869302
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; TITLE OF INVENTION: Reduced Protease Activity
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRES:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1B7

	Query Match	49.8%;	Score 607.5;	DB 2;	Length 228;
	Best Local Similarity	55.6%;	Pred. No. 1.1e-51;		
	Matches 110;	Conservative 36;	Mismatches 49;	Indels 3;	Gaps 3;
QY	34	IVNGENAVLGSPWQVQVSLQDSSGFHFCGGSLISQSWVWVTAACHNCVSPGRHFWVLGEYDRS	93		
Db	1	IVNGEAVPGSPWQVQVSLQDKTGFHFCGGSLINENWVWVTAARCGVTTs-DVVAGEFDQG	59		
QY	94	SNAPLQVLVSRAIHPSPNSNTMNDVILLKLASPAQVTTIRISPVCLASSNEALTEGL	153		
		: : : : : : : : : : : : : : : :			
Db	60	SSSEKIQKLIAKVFVKNKSNYSUTINNDITLLKSTAAFSQTVSAVCLVCSASDDFAAGT	119		

QY	154	TCVTTGRLSGVGNVTPAHLQVALPLVTVNOCROYWDSITDSMICAGGASSCOGD	213
		: : : : :	
Db	120	TCVTTGWL-TRYAN-TPDRLQQAQLPILLSNTNKKYWGTKIDAMICAGASGVSSCMGD	177
		: : : : :	
QY	214	SGGPLVCQKGNWTWVLGI	231
		: : : : :	
Db	178	SGGPLVCKKNGAWTLVGI	195
		: : : : :	

```

RESULT 10
US-08-482-816-10
; Sequence 10, Application US/08482816
; Patent No. 5935573
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; Zip: MSG lR7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482.816
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-494 MIS.vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; PS-08-482-816-10

```

	Query Match	49.8%;	Score 607.5;	DB 2;	Length 228;
	Best Local Similarity	55.6%;	Pred. No. 1.le-51;		
	Matches 110;	Conservative 36;	Mismatches 49;	Indels 3;	Gaps 3;
QY	34	I V G E N A V L G S W P Q V S L O D S S G F H C G G S L I S Q S W Y V T A A H C N V S P R H F V V L G E Y D R S	93		
		: : :			
D b	1	I V G E A V P C S W P Q V S L O D K T G F H C G G S L I N E N V V T A A H C G V T T S - D V V Y A G E D Q G	59		
		: : :			
QY	94	S N A E P L Q V L S V S R A I T H P S W N S T M N D V T L L K L A G P A Q Y T T R I S P V C L A S S N E A L T E G L	153		
		: : : : : : : : : : :			
D b	60	S S E K T Q K L A I A V F N K S Y N S L T I N N D I T L L K L S T A A S F Q V S A V C L P S A S D D F A A G T	119		
		: : : : : : : : : : :			
QY	154	T C V T T G W G R L S G V G N V T P A H L Q V A L P L V T V N C R Q Y W D S S I T D S M I C A G G A G A S C Q G D	213		
		: : : : : : : : : : :			

RESULT 14
US-09-074-660-10
: Sequence 10, Application US/09074660
: Patent No. 6020183
: GENERAL INFORMATION:
: APPLICANT: LOOSMORE, Sheena M.
: APPLICANT: YANG, Yan-Ping
: APPLICANT: CHONG, Pele
: APPLICANT: COOMEN, Raymond P.

RESULT 15
US-09-074-659-10
: Sequence 10, Application US/09074659
: Patent No. 6025342
: GENERAL INFORMATION:
: APPLICANT: LOOSMORE, Sheena M.
: APPLICANT: YANG, Yan-ping
: APPLICANT: CHONG, Pele

Search completed: December 20, 2002, 15:20:36
Job time : 14.087 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:14:08 ; Search time 28.6087 Seconds
(without alignments)
1663.721 Million cell updates/sec

Title: US-09-856-319B-4_COPY_34_264

Perfect score: 1252

Sequence: 1 IVNGENAVPGSWPQVSLQD.....AMVTRVSKFSTWVQVMAYN 231

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1252	100.0	264	11 Q9ER05	Q9ER05 mus musculus
2	1248	99.7	264	11 Q9D7P8	Q9D7P8 mus musculus
3	1242	99.2	264	11 Q9D960	Q9D960 mus musculus
4	1199	95.8	264	11 Q9EQ28	Q9EQ28 rattus norv
5	841.5	67.2	261	13 Q9W704	Q9W7q4 paralichthy
6	749.5	59.9	260	13 Q9W7Q3	Q9W7q3 paralichthy
7	727.5	58.1	263	11 Q9DC86	Q9DC86 mus musculus
8	725.5	57.9	263	11 Q9CR35	Q9CR35 mus musculus
9	719.5	57.5	263	11 Q9D8X8	Q9D8x8 mus musculus
10	712.5	56.9	263	13 Q9PWQ6	Q9PWq6 gadus morhu
11	674	53.8	164	11 Q9DC82	Q9DC82 mus musculus
12	511.5	40.9	269	11 Q9D7F9	Q9D7F9 mus musculus
13	511.5	40.9	269	11 Q9CQ52	Q9CQ52 mus musculus
14	496	39.6	266	13 Q92077	Q92077 gadus morhu
15	494.5	39.5	269	6 Q95KW7	Q95kw7 bos taurus
16	493	39.4	249	13 Q9W7Q1	Q9W7q1 paralichthy

17	487	38.9	1524	13	Q91674	Q91674 xenopus lae
18	486.5	38.9	260	13	Q9W7P9	Q9W7p9 paralichthy
19	485.5	38.8	267	5	Q9BK47	Q9BK47 luidia foli
20	465.5	37.2	810	4	Q15146	Q15146 homo sapien
21	465	37.1	812	11	Q91WJ5	Q91Wj5 mus musculu
22	462.5	36.9	266	13	Q9W7Q0	Q9W7q0 paralichthy
23	460	36.7	268	13	Q9W7Q2	Q9W7q2 paralichthy
24	459.5	36.7	270	4	Q96Q18	Q96ql8 homo sapien
25	457.5	36.5	257	6	O19023	O19023 macaca mula
26	457.5	36.5	277	5	O96899	O96899 scolopendra
27	453	36.2	269	4	Q96QV5	Q96qv5 homo sapien
28	451.5	36.1	339	11	Q99144	Q99144 mus musculu
29	448.5	35.8	273	11	Q921N4	Q921n4 mus musculu
30	444.5	35.5	334	6	O46507	O46507 papio hanad
31	444	35.5	812	11	Q9R0W3	Q9R0w3 rattus norv
32	442.5	35.3	270	13	Q91039	Q91039 gadus morhu
33	442.5	35.3	453	11	Q8VDE0	Q8Vde0 mus musculu
34	439.5	35.1	244	13	Q8QGW3	Q8qgw3 angullia ja
35	438.5	35.0	266	6	O46644	O46644 macaca fasc
36	438.5	35.0	331	11	Q8RIA6	Q8ria6 mus musculu
37	438	35.0	246	11	Q9QK9	Q9qk9 mus musculu
38	437.5	34.9	266	11	Q91X79	Q91x79 mus musculu
39	436.5	34.9	247	13	O42608	O42608 petromyzon
40	435.5	34.8	266	11	Q9D936	Q9D936 mus musculu
41	435.5	34.8	266	11	Q9W7Q6	Q9W7q6 paralichthy
42	435	34.7	238	13	Q921R9	Q921r9 mus musculu
43	435	34.7	246	11	Q921R9	Q921r9 mus musculu
44	434	34.7	246	11	Q9R0T7	Q9R0t7 mus musculu
45	432.5	34.5	806	6	O18783	O18783 macropus eu

ALIGNMENTS

RESULT 1

Q9ER05
ID Q9ER05 PRELIMINARY; PRT; 264 AA.
AC Q9ER05;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Chymopasin (Chymotrypsin A CTRA-1).
GN CTRL OR CTRAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Mitsui S., Yamaguchi N.;
RT "Molecular cloning of mouse chymopasin."
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129S6/SVEVTAC; TISSUE=SPLEEN;
RA Bjoernslett M.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.
DR EMBL; AB016228; BAB20275.1; -
DR EMBL; AF236365; AAL11034.1; -
DR HSSP; P00766; 4CHA.
DR MEROPS; S01.256; -
DR MGD; MGI:88558; Ctrl.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.

```

SQ SEQUENCE 264 AA; 28135 MW; 1D979709A07056C2 CRC64;
Query Match 100.0%; Score 1252; DB 11; Length 264;
Best Local Similarity 100.0%; Pred. No. 9.7e-107;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IVNGENAVPGSPWQVSLQDNTGFHFCGSLISPNVNTAAHCQVTPGRHFVVLGEYDRS 60
DB 34 IVNGENAVPGSPWQVSLQDNTGFHFCGSLISPNVNTAAHCQVTPGRHFVVLGEYDRS 93
QY 61 SNAEPQVLSIARAIHPNNNANTMNDLTLLKLSPARYTAQVSPVCLASTNEALPSGL 120
DB 94 SNAEPQVLSIARAIHPNNNANTMNDLTLLKLSPARYTAQVSPVCLASTNEALPSGL 153
QY 121 TCVTGWRISGVNTPARLQOVLPVLTVNOCROYWGARITDAMICAGGSCSCQGD 180
DB 154 TCVTGWRISGVNTPARLQOVLPVLTVNOCROYWGARITDAMICAGGSCSCQGD 213
QY 181 SGGPLVCQKGTWVLIGIVSGTKNCNIQAPAMYTRVSKFSTWVNOVMAYN 231
DB 214 SGGPLVCQKGTWVLIGIVSGTKNCNIQAPAMYTRVSKFSTWVNOVMAYN 264

RESULT 2
Q9D7P8 PRELIMINARY; PRT; 264 AA.
ID Q9D7P8 AC Q9D960 PRELIMINARY; PRT; 264 AA.
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DE 1810004D15RIK protein.
GN CTRL OR 1810004D15RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa K., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Nardone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AK009019; BAB26029.1; -.
DR HSSP; P00766; 4CHA.
DR MEROPS; S01.256; -.
DR MGD; MGI:88558; Ctrl.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; Chymotrypsin.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28167 MW; 1D979469A07056C2 CRC64;
Query Match 99.7%; Score 1248; DB 11; Length 264;
Best Local Similarity 99.6%; Pred. No. 2.3e-106;
Matches 230; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IVNGENAVPGSPWQVSLQDNTGFHFCGSLISPNVNTAAHCQVTPGRHFVVLGEYDRS 60
DB 34 IVNGENAVPGSPWQVSLQDNTGFHFCGSLISPNVNTAAHCQVTPGRHFVVLGEYDRS 93
QY 61 SNAEPQVLSIARAIHPNNNANTMNDLTLLKLSPARYTAQVSPVCLASTNEALPSGL 120
DB 94 SNAEPQVLSIARAIHPNNNANTMNDLTLLKLSPARYTAQVSPVCLASTNEALPSGL 153
QY 121 TCVTGWRISGVNTPARLQOVLPVLTVNOCROYWGARITDAMICAGGSCSCQGD 180
DB 154 TCVTGWRISGVNTPARLQOVLPVLTVNOCROYWGARITDAMICAGGSCSCQGD 213
QY 181 SGGPLVCQKGTWVLIGIVSGTKNCNIQAPAMYTRVSKFSTWVNOVMAYN 231
DB 214 SGGPLVCQKGTWVLIGIVSGTKNCNIQAPAMYTRVSKFSTWVNOVMAYN 264

RESULT 3
Q9D960 PRELIMINARY; PRT; 264 AA.
ID Q9D960 AC Q9D960 PRELIMINARY; PRT; 264 AA.
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DE 1810004D15RIK protein.
GN CTRL OR 1810004D15RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa K., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Nardone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AK007333; BAB24967.1; -.
DR HSSP; P00766; 4CHA.
DR MEROPS; S01.256; -.
DR MGD; MGI:88558; Ctrl.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; Chymotrypsin.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28167 MW; 1D979469A07056C2 CRC64;
Query Match 99.7%; Score 1248; DB 11; Length 264;
Best Local Similarity 99.6%; Pred. No. 2.3e-106;
Matches 230; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IVNGENAVPGSPWQVSLQDNTGFHFCGSLISPNVNTAAHCQVTPGRHFVVLGEYDRS 60
DB 34 IVNGENAVPGSPWQVSLQDNTGFHFCGSLISPNVNTAAHCQVTPGRHFVVLGEYDRS 93
QY 61 SNAEPQVLSIARAIHPNNNANTMNDLTLLKLSPARYTAQVSPVCLASTNEALPSGL 120
DB 94 SNAEPQVLSIARAIHPNNNANTMNDLTLLKLSPARYTAQVSPVCLASTNEALPSGL 153
QY 121 TCVTGWRISGVNTPARLQOVLPVLTVNOCROYWGARITDAMICAGGSCSCQGD 180
DB 154 TCVTGWRISGVNTPARLQOVLPVLTVNOCROYWGARITDAMICAGGSCSCQGD 213
QY 181 SGGPLVCQKGTWVLIGIVSGTKNCNIQAPAMYTRVSKFSTWVNOVMAYN 231
DB 214 SGGPLVCQKGTWVLIGIVSGTKNCNIQAPAMYTRVSKFSTWVNOVMAYN 264
```



```

DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28151 MW; 1D979719E07C16DE CRC64;

Query Match
Best Local Similarity 99.2%; Score 1242; DB 11; Length 264;
Matches 230; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IVNGENAVPGSWPQVSLQDNTGFHFCGGLISPNWVYTAHCKVTPGRHFVILGEYDRS 60
DB 34 IVNGENAVPGSWPQVSLQDNTGFHFCGGLISPNWVYTAHCKVTPGRHFVILGEYDRS 93
QY 61 SNAEPVQVLSARAIHPNNWNTMNDLTLKLSAPRYTAQVSPVCLASTNEALPSGL 120
DB 94 SNAEPVQVLSARAIHPNNWNTMNDLTLKLSAPRYTAQVSLVCLASTNEALPSGL 153
QY 121 TCVTTGWRISGVGNVTTPARLQVVLPLVTYNQCRQYWGARITDAMICAGSGASSCGD 180
DB 154 TCVTTGWRISGVGNVTTPARLQVVLPLVTYNQCRQYWGARITDAMICAGSGASSCGD 213
QY 181 SGGPLVCOKGNTWVLIGIVSGTKNCNIQAPAMTRYSKFSTWYNQVAYN 231
DB 214 SGGPLVCOKGNTWVLIGIVSGTKNCNIQAPAMTRYSKFSTWYNQVAYN 264

RESULT 4
Q9EQZ8 PRELIMINARY; PRT; 264 AA.
AC Q9EQZ8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-JUN-2002 (TrEMBLrel. 16, Last sequence update)
DE Chymopasin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-RAT PANCREAS;
RA Sogame Y., Mitsui S., Kataoka K., Kashima K., Kato M., Sakagami J.,
RA Yamaguchi N.;
RT "Molecular cloning of rat chymopasin.";
RT Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB020757; BAB20287.1; -.
DR HSSP; P00766; 4CHA.
DR MEROPS; S01.256; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28116 MW; F9ED5D210FD3500E CRC64;

Query Match
Best Local Similarity 95.8%; Score 1199; DB 11; Length 264;
Matches 214; Conservative 16; Mismatches 1; Indels 0; Gaps 0;

QY 1 IVNGENAVPGSWPQVSLQDNTGFHFCGGLISPNWVYTAHCKVTPGRHFVILGEYDRS 60
DB 34 IVNGENAVPGSWPQVSLQDNTGFHFCGGLISPNWVYTAHCKVTPGRHFVILGEYDRS 93
QY 61 SNAEPVQVLSARAIHPNNWNTMNDLTLKLSAPRYTAQVSPVCLASTNEALPSGL 120
DB 94 SNAEPVQVLSARAIHPNNWNTMNDLTLKLSAPRYTAQVSPVCLASTNEALPSGL 153

```

```

QY 121 TCVTTGWRISGVGNVTTPARLQVVLPLVTYNQCRQYWGARITDAMICAGSGASSCGD 180
DB 154 TCVTTGWRISGVGNVTTPARLQVVLPLVTYNQCRQYWGARITDAMICAGSGASSCGD 213
QY 181 SGGPLVCOKGNTWVLIGIVSGTKNCNIQAPAMTRYSKFSTWYNQVAYN 231
DB 214 SGGPLVCOKGNTWVLIGIVSGTKNCNIQAPAMTRYSKFSTWYNQVAYN 264

RESULT 5
Q9W7Q4 PRELIMINARY; PRT; 261 AA.
AC Q9W7Q4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chymotrypsinogen 1.
OS Paralicthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidae; Paralicthyidae; Paralicthys.
OX NCBI_TaxID=8255;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PANCREAS;
RA Suzuki T., Srivastava A.S., Kurokawa T.;
RT "Japanese flounder mRNA for chymotrypsinogen 1.";
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB029753; BAA82365.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.256; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 261 AA; 28184 MW; D7090A9D65395B7D CRC64;

Query Match
Best Local Similarity 67.2%; Score 841.5; DB 13; Length 261;
Matches 153; Conservative 28; Mismatches 48; Indels 3; Gaps 2;

QY 1 IVNGENAVPGSWPQVSLQDNTGFHFCGGLISPNWVYTAHCKVTPGRHFVILGEYDRS 60
DB 32 IVNGENAVPGSWPQVSLQDNTGFHFCGGLISPNWVYTAHCKVTPGRHFVILGEYDRS 91
QY 61 SNAEPVQVLSARAIHPNNWNTMNDLTLKLSAPRYTAQVSPVCLASTNEALPSGL 120
DB 92 YNEPIQVMSIARAIHPNNWNTMNDLTLKLSAPRYTAQVSPVCLASTNEALPSGL 151
QY 121 TCVTTGWRISGVGNVTTPARLQVVLPLVTYNQCRQYWG-ARITDAMICAGSGASSCG 179
DB 152 KCVTTGWRGTQTS--SPRYLQOTSLPLSPAQCKQYNGYRITDAMICAGSGASSCG 209
QY 180 DSGPLVCOKGNTWVLIGIVSGTKNCNIQAPAMTRYSKFSTWYNQVAYN 231
DB 210 DSGPLVCOKGNTWVLIGIVSGTKNCNIQAPAMTRYSKFSTWYNQVAYN 261

RESULT 6
Q9W7Q3 PRELIMINARY; PRT; 260 AA.
AC Q9W7Q3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

```

01-WAR-2002 (TREMELrel. 20, Last annotation update)
 DE Chymotrypsinogen 2.
 OS Paralicthys olivaceus (Flounder).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
 OC Pleuronectoidae; Paralicthidae; Paralicthidae; Paralicthys.
 OX NCBI_TaxID=8255;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PANCREAS;
 RA Suzuki T., Srivastava A.S., Kurokawa T.;
 RT "Japanese flounder mRNA for chymotrypsinogen 2.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 TRYPsin FAMILY.
 CC EMBL: AB029754; BAA82366.1; -.
 DR HSSP: P00766; ICHG.
 DR MEROPS: S01.152; -.
 DR InterPro: IPR001314; Chymotrypsin.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; tryp-Spc; 1.
 DR PROSITE: PS0240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 260 AA; 27793 MW; 9F583044E22F78C0 CRC64;
 Query Match 59.9%; Score 749.5; DB 13; Length 260;
 Best Local Similarity 57.6%; Pred. No. 9.8e-61;
 Matches 133; Conservative 35; Mismatches 62; Indels 1; Gaps 1;
 QY 1 IVNGENAVPGSPWQVSLQDNTGFHFCGSLSPNVTAAHCQVTPGRHFVVLGEYDRS 60
 Db 31 IVNGEALPHSPWQVSLQDNTGFHFCGSLSPNVTAAHCNVRTS-HRVILGEHDS 89
 QY 61 SNAEPQVLSIARAIHPNNANTNMNDLTLKLPASARYTAQVSPVCLASTNEALPSGL 120
 Db 90 SNAEDIQVVKVGVKPRYNGTYTINDILLIKLAAPQNMHRVSPVCVAETSDNFAGM 149
 QY 121 TCVTGWGRISGVNTPARLQVVLPLVTNQCQRYGARTDAMICAGSGSCQGD 180
 Db 150 KCVTSWGLTRHNPADPTPALQQAALPLTNDCCRYGKNSLNICAGSAGSCSCMD 209
 QY 181 SGGPLVCQKGNVTWLVIGVSWGTKNICQAPAMYTRVSKFSTWVNOYMAYN 231
 Db 210 SGGPLVCQKAGAWTLGVISWGSCTPTPTMGYARVTELRAWMDQTIANN 260
 RESULT 7
 Q9DC86 PRELIMINARY; PRT; 263 AA.
 ID Q9DC86
 AC Q9DC86;
 DT 01-JUN-2001 (TREMELrel. 17, Created)
 DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)
 DE 2200008D09rik protein.
 GN 2200008D09rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=SPLEEN;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boilelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 403:685-690(2001).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 TRYPsin FAMILY.
 CC EMBL: AK003060; BAB22539.1; -.
 DR HSSP: P00766; IGCT.
 DR MEROPS: S01.152; -.
 DR MGD: MGI:1913723; 2200008D09rik.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; Tryp_Spc; 1.
 DR PROSITE: PS0240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 263 AA; 27821 MW; 2620A27AFBA5D04D CRC64;
 Query Match 58.1%; Score 727.5; DB 11; Length 263;
 Best Local Similarity 56.3%; Pred. No. 1e-58;
 Matches 130; Conservative 37; Mismatches 63; Indels 1; Gaps 1;
 QY 1 IVNGENAVPGSPWQVSLQDNTGFHFCGSLSPNVTAAHCQVTPGRHFVVLGEYDRS 60
 Db 34 IVNGEDAIPGSPWQVSLQDRTGFHFCGSLSPNVTAAHCQVKT-TNVVAGEFDQG 92
 QY 61 SNAEPQVLSIARAIHPNNANTNMNDLTLKLPASARYTAQVSPVCLASTNEALPSGL 120
 Db 93 SDEENQVQLKIAQVKNPKFNFTNRNDLTLKLPATPAQFSETSAVCLFTVDDDFPAGT 152
 QY 121 TCVTGWGRISGVNTPARLQVVLPLVTNQCQRYGARTDAMICAGSGSCQGD 180
 Db 153 LCATTGWGKTKYNALKTDPKLOQALPIVSEAKCKESGSKITDVMICAGASVSSCMGD 212
 QY 181 SGGPLVCQKGNVTWLVIGVSWGTKNICQAPAMYTRVSKFSTWVNOYMAYN 231
 Db 213 SGGPLVCQKDGWTLGVISWGSCTSTPAVYARVATLMPWQVLEAFN 263
 RESULT 8
 Q9CR35 PRELIMINARY; PRT; 263 AA.
 ID Q9CR35
 AC Q9CR35;
 DT 01-JUN-2001 (TREMELrel. 17, Created)
 DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)
 DE 2200008D09rik protein.
 GN 2200008D09rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=STOMACH, SPLEEN, AND PANCREAS;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

```

RN      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX      MEDLINE=21085660; PubMed=11217851;
RA      Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA      Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA      Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA      Saito T., Okazaki Y., Gotojori T., Bono H., Kasukawa T., Saito R.,
RA      Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA      Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA      Kuehl P., Lewis S., Matsuo Y., Mikaido I., Pesole G., Quackenbush J.,
RA      Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA      Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA      Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA      Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA      Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA      Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA      Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA      Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA      Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA      Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA      Hayashizaki Y.;
RT      "Functional annotation of a full-length mouse cDNA collection.";
RL      Nature 409:685-690(2001)
CC      - 1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC      TRYPSIN FAMILY.
DR      EMBL; AK007566; BAB25112.1; -.
DR      HSSP; P00766; LGCT.
DR      MEROPS; S01.152; -.
DR      MGD; MGI:1913773; 2200008D09Rik.
DR      InterPro; IPR001334; Chymotrypsin.
DR      InterPro; IPR001254; Ser.protease.Try.
DR      Pfam; PF00089; trypsin; 1.
DR      PRINTS; PR00722; CHYMOTRYPSIN.
DR      SMART; SM00020; Tryp.SPC; 1.
DR      PROSITE; PS02040; TRYPSIN_DOM; 1.
DR      PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR      PROSITE; PS00135; TRYPSIN_SER; 1.
KW      Hydrolase. Serine protease.
SQ      SEQUENCE 263 AA; 27898 MW; C0638FB8F905A92F CRC64;

Query Match          57.5%; Score 719.5; DB 11; Length 263;
Best Local Similarity 55.8%; Pred.No.5.5e-58;
Matches 129; Conservative 36; Mismatches 65; Indels 1; Gaps

QY      1 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNVVYTAAHQCVTPGRHFVVLGYDRS 60
DB      IVNGEIDAIPGSGWPQVSLQDRTGFHFCGGYLIISNNVVYTAAHGKVT-TDVVVAGEFDQ 92
QY      61 SNAEPQVVLSTARAITHPNKNANTMNDLLKLLASPARYTAQVSPVCLASTNEALPSGL 120
DB      SDEENQVQLKIAQVFNKPFNSFTVRNDITLLKLTAPQSETSVAVCLPTVDDDPAGT 152
QY      121 TCVTTWGRISGVGNVTPARLQVVLPLVTVNCQVWGARITDAMICAGSGASSCGQD 180
DB      LCATTGKGKTKYNALKTPDKLQQAALPIVSEAKCKESWGSKIITDVMICAGASGVSSCMGD 212
QY      181 SGGPLVCQKGNVTWVLIGVSWGTKNCNIQAPAMATRVYSKFSTWVNWVAYN 231
DB      SGGPLVCQKGVNVTWVLIGVSWGTFSGFSTSTPVAIVARTALMPVQETLEIAN 263

```

RESULT 10					
Q9PWQ6					
ID	Q9PWQ6	PRELIMINARY;	PRT;	263 AA.	
AC	Q9PWQ6;				
DT	01-MAY-2000	(TrEMBLrel. 13, Created)			
DT	01-MAY-2000	(TrEMBLrel. 13, Last sequence update)			
DE	01-MAR-2002	(TrEMBLrel. 20, Last annotation update)			
DE	Chymotrypsin B precursor (EC 3.4.21.1).				
GN	CHYB.				
OS	Gadus morhua (Atlantic cod).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;				

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthopterygii; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
 OX NCBI_TaxID=8049;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PYLORIC CAECA;
 RA MEDLINE=20464334; PubMed=11011764;
 RA Spilliaert R., Gudmundsdottir A.;
 RA "Molecular Cloning of the Atlantic Cod Chymotrypsinogen B";
 RL Microb. Comp. Genomics 5:41-50(2000).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL; AJ242521; CAB43766.1; -.
 DR HSSP; P00766; ICHG.
 DR MEROPS; S01.152; -.
 DR InterPro: IPR001314; Chymotrypsin.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS0134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS0135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Signal.
 FT SIGNAL 1 16
 FT CHAIN 17 263 CHYMOTRYPSIN B.
 SQ SEQUENCE 263 AA; 28175 MW; EF61B18A34EE5E7C CRC64;
 Query Match 56.98; Score 712.5; DB 13; Length 263;
 Best Local Similarity 54.9; Pred. No. 2.4e-57;
 Matches 129; Conservative 41; Mismatches 58; Indels 7; Gaps 5;
 QY 1 IVNGENAVPGSWPQVSLQDNTGTFHFCGGLISPNVYTAACHCOVTPGRHFVVLGEYDRS 60
 DB 32 IVNGEAVPHSWPQVSLQDSNGFHCGLINENWVYTAACHVTR-YHRVIYGEHDKS 90
 QY 61 -SNAEPQVLSIARATHHPNNANTMNDLTLKLPARYTAQVSPVCLASTNEALPSG 119
 DB 91 RASDENIQLKPSMVFTHPKWDSRTINNDISLIKLPASVLTGTVNVPCLGESSDVFAPG 150
 QY 120 LTCVTTGNG--RISGVNVTARLQOVVPLVTVNOCROYWGR-ITDAMICAGSGASS 176
 DB 151 MKCVTSGLTRYNAPG--TPNKIQQAALPLMSNEECSTQWGNMISDMVICAGAAGATS 208
 QY 177 CGSDSGPLVCQKGTWTVLIGVSWGKNCNIQAPAMTRYVSKFSTWVNOVMAYN 231
 DB 209 CMGDSGGLVCQKQNVTVLIGVSWGSRCSVTTPAYARVYTELGRWVDQILAN 263
 RESULT 11
 ID Q9DC82 PRELIMINARY; PRT; 164 AA.
 AC Q9DC82;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE 0910001G08Rik protein.
 GN CTRL OR 0910001G08Rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=SPLEEN;
 RA MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection";
 RL Nature 409:685-690(2001).
 DR EMBL; AK003074; BAB22549.1; -.
 DR HSSP; P00763; IDPO.
 DR MEROPS; S01.997; -.
 DR MGD; MGI:88558; Ctrl.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS0135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 164 AA; 17707 MW; 83791FD829ABEDD6 CRC64;
 Query Match 53.8; Score 674; DB 11; Length 164;
 Best Local Similarity 82.2; Pred. No. 4.5e-54;
 Matches 129; Conservative 5; Mismatches 21; Indels 2; Gaps 2;
 QY 77 HPNNA-NTWNNDLTLKLPAS-ARYTAQVSPVCLASTNEALPSGLTCVTGWRISGVG 134
 DB 8 HPHWGGHPLKTRLFLEFFPHPKTKFPFGFLASTNEALPSGLTCVTGWRISGVG 67
 QY 135 NVTARLQOVVPLVTVNOCROYWGRITDAMICAGSGASSCGDSGGLVCQKGTWV 194
 DB 68 NVTARLQOVVPLVTVNOCROYWGRITDAMICAGSGASSCGDSGGLVCQKGTWV 127
 QY 195 LIGVSWGKNCNIQAPAMTRYVSKFSTWVNOVMAYN 231
 DB 128 LIGVSWGKNCNIQAPAMTRYVSKFSTWVNOVMAYN 164
 RESULT 12
 ID Q9D7T9 PRELIMINARY; PRT; 269 AA.
 AC Q9D7T9;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE 2310074F01Rik protein.
 GN ELA3B OR 2310074F01Rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=STOMACH;
 RA MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Wyszaw-Boris A., Yoshida K., Weitz C., Whittaker C., Wilming L.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL; AK008858; BAB25932.1; -;
 DR HSSP; P05805; 1FON.
 DR MEROPS; S01.154; -;
 DR MGD; MGI:1915118; Ela3b.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 269 AA; 28977 MW; 9F43F769DDB2A7CF CRC64;
 Query Match 40.9%; Score 511.5; DB 11; Length 269;
 Best Local Similarity 44.0%; Pred. No. 6.3e-39;
 Matches 107; Conservative 37; Mismatches 86; Indels 13; Gaps 8;
 QY 1 IVNGENAVPGSPWQVLSQ---DNTGFHFCGSLSPNNVYTAHCQVTPGRHFVVLGEY 57
 DB 28 VYNGEAVPHSPWQVLSQYKDGSHHTCGSLTPDMLTAGHCISTRTYQVVLGEH 87
 QY 58 DRSSNAEPVQVLSI--ARAITHPNNTMN--NDLTLLKLASPARYAQVSPVCLASTN 113
 DB 88 ERGVEEGEQVTPINAGDLFVHPKNSMCVSCGNDIALVKLSRAQLGDAVQLACLPPAG 147
 QY 114 EALPSGLTCVTGWGRISGVGNVTPARLQVVLPLVTVNCQKQ--YWGARITDAMICAGG 171
 DB 148 EILPAGPCYISGWRSLSTNGPL-PDKLQALLPVVDYEHCSRNNWGLSVKTTMYCAGG 206
 QY 172 SGASSCGDSGGLPVCQKGN-TWVLIGIVSW-GTKNCN-IQAPAMYTRVSKFSTWVNOVM 228
 DB 207 DIQSGCNGSDGGLNCPADNGTQVHGVTSEVSSLGNCNLTFRKPTVTRVSAFIDWIEETI 266
 QY 229 AVN 231
 DB 267 ANN 269
 RESULT 13
 ID Q9CQ52 PRELIMINARY; PRT; 269 AA.
 AC Q9CQ52;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE 2310074F01Rik protein.
 GN ELA3B OR 2310074F01Rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=TONGUE;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King C., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baidarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustcich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wyszaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL; AK010149; BAB26734.1; -;
 DR HSSP; AK009129; BAB26092.1; -;
 DR HSSP; P05805; 1FON.
 DR MEROPS; S01.154; -;
 DR MGD; MGI:1915118; Ela3b.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 269 AA; 28904 MW; C543F76957B2A7CE CRC64;
 Query Match 40.9%; Score 511.5; DB 11; Length 269;
 Best Local Similarity 44.0%; Pred. No. 6.3e-39;
 Matches 107; Conservative 37; Mismatches 86; Indels 13; Gaps 8;
 QY 1 IVNGENAVPGSPWQVLSQ---DNTGFHFCGSLSPNNVYTAHCQVTPGRHFVVLGEY 57
 DB 28 VYNGEAVPHSPWQVLSQYKDGSHHTCGSLTPDMLTAGHCISTRTYQVVLGEH 87
 QY 58 DRSSNAEPVQVLSI--ARAITHPNNTMN--NDLTLLKLASPARYAQVSPVCLASTN 113
 DB 88 ERGVEEGEQVTPINAGDLFVHPKNSMCVSCGNDIALVKLSRAQLGDAVQLACLPPAG 147
 QY 114 EALPSGLTCVTGWGRISGVGNVTPARLQVVLPLVTVNCQKQ--YWGARITDAMICAGG 171
 DB 148 EILPAGPCYISGWRSLSTNGPL-PDKLQALLPVVDYEHCSRNNWGLSVKTTMYCAGG 206
 QY 172 SGASSCGDSGGLPVCQKGN-TWVLIGIVSW-GTKNCN-IQAPAMYTRVSKFSTWVNOVM 228
 DB 207 DIQSGCNGSDGGLNCPADNGTQVHGVTSEVSSLGNCNLTFRKPTVTRVSAFIDWIEETI 266
 QY 229 AVN 231
 DB 267 ANN 269
 RESULT 14
 ID Q92077 PRELIMINARY; PRT; 266 AA.
 AC Q92077;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Elastase precursor (EC 3.4.21.37) (Leukocyte elastase) (Lysosomal
 DE elastase) (Neutrophil elastase) (Bone marrow serine protease)
 DE (MEDULLASIN).
 GN ELB.
 OS Gadus morhua (Atlantic cod).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
 OX NCBI_TaxID=8049;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE-96262679; PubMed-8925447;
 RA Gudmundsdottir E., Spilliaert R., Yang Q., Craik C.S., Bjarnason J.B.,
 RA "Isolation and characterization of two cDNAs from Atlantic cod
 RT encoding two distinct psychrophilic elastases.";
 RL Comp. Biochem. Physiol. 113B:795-801(1996).
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF PROTEINS, INCLUDING ELASTIN.
 CC PREFERENTIAL CLEAVAGE: VAL-I-XAA > ALA-I-XAA.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL: U56936; AAB38350.1; -.
 DR HSSP: P05805; IFON.
 DR MEROPS: S01.155; -.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser.protease_Try.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; TRYPSIN_DOM; 1.
 DR PROSITE: PS0240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Signal; zymogen.
 FT SIGNAL 1 15 POTENTIAL.
 FT PROPEP 16 27 ACTIVATION PEPTIDE (POTENTIAL).
 FT CHAIN 28 266 ELASTASE.
 SQ SEQUENCE 266 AA; 28533 MW; B786B52C7159E2E CRC64;

 Query Match 39.6%; Score 496; DB 13; Length 266;
 Best Local Similarity 42.4%; Pred. No. 1.6e-37;
 Matches 101; Conservative 48; Mismatches 77; Indels 12; Gaps 9;

QY 1 IVNGENAVPCSWPQVSLDNTG---FHFCGSLISPNWVVTAAHCQVTPGRHFVVLGEY 57
 DB 28 VVGEDVRVHSWFQASLQKSNSTHYTCGGTLLAPQWVWTAHC-IGSTRYRVLLGKH 86
 QY 58 D-RSSNAEPVQVLSIARAIHPNNANTMNDTLKLSAPARYTAQVSPVCLASTNEAL 116
 DB 87 NMQDYNEAGSLAISPAAKIIHVKWDSRRINDIALIKLASPDVDSAIITPACVPDAEVL 146
 QY 117 PSLGLTCVTTGWGRISGVNVTAPLQOVVLPVTVNQCROY--WGARITDAMICAGSGA 174
 DB 147 ANGAPCVTGWRLTGGPTJADA-LQALLPVVDHACRYDWMGSLVTTSMYCAGGDGV 205
 QY 175 -SSCGDSDGGLVPCQKGN-TWVLIGIVSWGCK-NCNI-QAPAMYTRVSKFSTWINQVM 228
 DB 206 LASCNGDGGPLNQNADGSDVHGVSFGSSMGCNPKPSVTRVSAYIPINNVM 263

RESULT 15

Q95KW7 PRELIMINARY; PRT; 269 AA.
 AC Q95KW7;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Proproteinase E (Fragment).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PANCREAS;
 RA Seddi R., Guo X.-J., Chaix J.-C., Puigserver A.;
 RT "Nucleotide sequence of a bovine pancreatic proproteinase E cDNA."
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY057840; AAL23697.1; -.
 DR InterPro: IPR001254; Ser.protease_Try.
 DR Pfam: PF00089; trypsin; 1.
 DR PROSITE: PS0240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPSIN_SER; UNKNOWN_1.

KW Hydrolase; Serine protease.
 FT NON_TER 1
 SQ SEQUENCE 269 AA; 29010 MW; 352E4202B04B8B2D CRC64;

 Query Match 39.5%; Score 494.5; DB 6; Length 269;
 Best Local Similarity 43.2%; Pred. No. 2.3e-37;
 Matches 105; Conservative 37; Mismatches 88; Indels 13; Gaps 8;

QY 1 IVNGENAVPCSWPQVSLQ---DNTGFHFCGSLISPNWVVTAAHCQVTPGRHFVVLGEY 57
 DB 28 VVNGEDAVPYSNWSQVSLQEKDGFHHTCGSLIAPDWVVTAGHCISTSTRTYOVVLGEY 87
 QY 58 DRS--SNAEPVQVLSIARAIHPNNANTM--NNDTLKLSAPARYTAQVSPVCLASTN 113
 DB 88 DRSVLESGSEQVIPINAGDLFVHPLNNSNCVACGNDIALYKLSAQLDGKQVLANLPPAG 147
 QY 114 EALPSGLTCVTTGWGRISGVNVTAPLQOVVLPVTVNQCRO--YWGARITDAMICAGG 171
 DB 148 DILPNEAPCYISGGRLY-TGGPLPKLQKALLPVVDYEHCSQWDMWGITVKKTMVCAGG 206
 QY 172 SGASSCGDSDGGLVPCQKGN-TWVLIGIVSW-GTKNCN-IQAPAMYTRVSKFSTWINQVM 228
 DB 207 DTRSGCNGSDGGPLNCPAADGSMQVHGVTFSVAFGCGNTIKKPTVTRVSATIDWIDETI 266
 QY 229 AYN 231
 DB 267 ASN 269

Search completed: December 20, 2002, 15:18:57
 Job time : 29.6087 secs

Result 3

GenCore version 5.1.3

Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:12:13 ; Search time 35.6087 Seconds
(without alignments)
864.421 Million cell updates/sec

Title: US-09-856-319B-2_COPY_1_231

Perfect score: 1221

Sequence: 1 MLLSLTSLVLLGSSWGCG.....GDSGGPLVCQKGNWTWVLGI 231

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1221	100.0	264	21 AAB11710	Human serine prote
2	1093	89.5	264	21 AAB11711	Mouse serine prote
3	996	81.6	192	21 AAB34077	Human pancreatic c
4	684.5	56.1	263	23 AAU82738	Amino acid sequenc
5	657.5	53.8	245	21 AAY99596	Bovine chymotrypsi
6	621.5	50.9	231	22 AAB98504	Human chymotrypsin
7	472	38.7	269	7 AAP61724	Porcine elastase I
8	472	38.7	269	8 AAP70758	Pig pancreas elast
9	446.5	36.6	253	13 AAR29621	Porcine pancreatic
10	445	36.4	269	7 AAP60062	Sequence Of human

11	439.5	36.0	269	7 AAP61723	Human elastase II.
12	427.5	35.0	269	21 AAB54124	Human pancreatic c
13	427	35.0	269	8 AAP70760	Human pancreas ela
14	427	35.0	1052	22 AAM78338	Human protein seq
15	424	34.7	319	21 AAB36481	Fusion gene with h
16	424	34.7	319	22 AAB67541	Anino acid sequenc
17	422	34.6	252	7 AAP60058	Sequence of human
18	420.5	34.4	270	8 AAP70759	Human pancreas ela
19	420.5	34.4	270	23 AAU87691	Human pancreatic t
20	420.5	34.4	273	21 AAB54243	Human pancreatic c
21	418.5	34.3	253	7 AAP60059	Sequence of human
22	417	34.2	279	22 ABG20513	Novel human diagno
23	416.5	34.1	242	7 AAP60061	Sequence of human
24	414.5	33.9	269	7 AAP61076	Human elastase I.
25	410.5	33.6	343	23 AAU78547	/ Human prostatic pr
26	410.5	33.6	343	23 ABB07285	Anino acid sequenc
27	410.5	33.6	453	21 AAB44250	Human PRO382 (UNQ3
28	408.5	33.5	452	20 AAY41694	Human PRO382 (prote
29	408.5	33.5	453	22 AAU29055	Human PRO polypept
30	408.5	33.5	453	22 AAE06935	Human membrane-ty
31	408.5	33.5	453	23 AAE33020	Human trypsin fami
32	408.5	33.5	454	21 AAB32246	Tumour associated
33	408.5	33.5	454	22 AAU68911	Human TAGD-12, pro
34	408	33.4	248	21 AAB43572	Human cancer assoc
35	408	33.4	327	21 AAY72093	Human serine prote
36	408	33.4	327	23 AAE17921	Human gene 3 encod
37	408	33.4	394	23 ABB41994	Human ovarian anti
38	408	33.4	454	23 AAE33024	Human trypsin fami
39	408	33.4	454	23 AAU82745	Anino acid sequenc
40	407.5	33.4	268	17 AAR90682	Rat caldesrin cont
41	403.5	33.0	811	12 AAR12933	Plasminogen mutain
42	403.5	33.0	811	12 AAR12945	Plasminogen mutain
43	401.5	32.9	811	12 AAR12939	Plasminogen mutain
44	401.5	32.9	811	12 AAR12943	Plasminogen mutain
45	399.5	32.7	811	12 AAR12944	Plasminogen mutain

ALIGNMENTS

RESULT 1

AAB11710

ID AAB11710 standard; Protein; 264 AA.

XX AAB11710;

XX 23-OCT-2000 (first entry)

XX Human serine protease BSSP5 (hbSSP5) SEQ ID NO:2.

XX BSSP5: serine protease; human; hbSSP5; mouse; mBSSP5; brain;
XX diagnostic marker; antibody; transgenic animal; Alzheimer's disease;
XX epilepsy; cancer; inflammation; infertility; pancreatitis;
XX prostatic hypertrophy.

XX Homo sapiens.

XX WO200031243-A1.

XX 02-JUN-2000.

XX 19-NOV-1999; 99WO-JP06473.

XX 20-NOV-1998; 98JP-0347806.

XX (FUSO) FUSO PHARM IND LTD.

XX Uemura H, Okui A, Kominami K, Yamaguchi N, Mitsui S;

XX WPI: 2000-400058/34.

XX N-PSDB; AAA61733.

XX Serine proteases BSSP5, useful in detecting homologs, mutants and

PT polymorphic variants as markers for diagnosis of e.g. Alzheimer's
PT disease, epilepsy, cancer and inflammation, using blood, urine,
XX pancreas or other tissues -
XX
XX Claim 1; Page 51-52; 70pp; Japanese.
XX
CC The invention relates to novel serine proteases designated BSSP5
CC (AAB11710-B11711), and to nucleic acids encoding them (AAA61733-A61734).
CC The invention also relates to vectors and transformatants comprising BSSP5
CC nucleic acids; transgenic animals in which the expression level of BSSP5
CC can be varied; and an mBSSP5 knockout mouse. The invention additionally
CC encompasses anti-BSSP5 antibodies and methods of production of such
CC antibodies, methods of BSSP5 detection using the antibodies, and the
CC use of BSSP5 proteins or fragments as diagnostic markers for certain
CC medical conditions, e.g., pancreatitis. A method for detecting
CC pancreatitis comprising measuring BSSP5 concentration in the blood or
CC urine, and a pancreatitis diagnostic agent containing an anti-BSSP5
CC antibody is also disclosed. Nucleotides encoding BSSP5 were initially
CC isolated in a human brain cDNA library using degenerate PCR primers
CC (AAA61744-A61745) based on conserved regions of serine proteases. The
CC BSSP5 serine proteases and nucleotides encoding them are useful in
CC detecting homologues, mutants and polymorphic variants in biological
CC samples (e.g., blood, urine, brain, prostate gland, placenta, testis,
CC pancreas and spleen) as diagnostic markers for conditions such as
CC Alzheimer's disease, epilepsy, cancer, inflammation, infertility,
CC pancreatitis and prostatic hypertrophy. Sequence AAB11710 represents
CC human BSSP5 (hBSSP5), and sequence AAB11711 represents murine BSSP5
CC (mBSSP5).
XX
XX Sequence 264 AA;
XX
XX Query Match 100.0%; Score 1221; DB 21; Length 264;
XX Best Local Similarity 100.0%; Pred. No. 1.1e-98;
XX Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLLSLTSLVLGSSWGGCIPAIKPAKPSQRIVNGENAVLGSWPQVSLQDSSGFHFC 60
DB 1 MLLSLTSLVLGSSWGGCIPAIKPAKPSQRIVNGENAVLGSWPQVSLQDSSGFHFC 60
QY 61 GGSLSISQSWVVTAAHCNVSPGRHFVVLGEYDRSSNAEPQLQVLSVRAITHPSWNSTMMN 120
DB 61 GGSLSISQSWVVTAAHCNVSPGRHFVVLGEYDRSSNAEPQLQVLSVRAITHPSWNSTMMN 120
QY 121 DVTLLKLASPAQYTTTRISPVCLASSNEALTEGLTCVTTGWGRLSGVGNVTPAHLQOVALP 180
DB 121 DVTLLKLASPAQYTTTRISPVCLASSNEALTEGLTCVTTGWGRLSGVGNVTPAHLQOVALP 180
QY 181 LVTVNOCROYWDSSTITDSMICAGGAGASSCGDGGPLVCQKGTWVLIGI 231
DB 181 LVTVNOCROYWDSSTITDSMICAGGAGASSCGDGGPLVCQKGTWVLIGI 231
RESULT 2
AAB11711
ID AAB11711 standard; Protein; 264 AA.
XX
XX AAB11711;
XX
XX 23-OCT-2000 (first entry)
XX
XX Mouse serine protease BSSP5 (mBSSP5) SEQ ID NO:4.
XX
XX BSSP5; serine protease; human; hBSSP5; mouse; mBSSP5; brain;
KW diagnostic marker; antibody; transgenic animal; Alzheimer's disease;
KW epilepsy; cancer; inflammation; infertility; pancreatitis;
KW prostatic hypertrophy.
XX
XX Mus sp.
XX
XX WO200031243-A1.
XX
XX 02-JUN-2000.
XX

PF 19-NOV-1999; 99WO-JP06473.
XX
PR 20-NOV-1998; 98JP-0347806.
XX
PA (FUSO) FUSO PHARM IND LTD.
XX
XX Uemura H, Okui A, Kominami K, Yamaguchi N, Mitsui S;
PI
XX WPI; 2000-400058/34.
DR N-PSDB; AAA61734.
DR
XX Serine proteases BSSP5, useful in detecting homologs, mutants and
PT polymorphic variants as markers for diagnosis of e.g. Alzheimer's
PT disease, epilepsy, cancer and inflammation, using blood, urine,
PT pancreas or other tissues -
XX
XX Claim 3; Page 55-56; 70pp; Japanese.
XX
XX The invention relates to novel serine proteases designated BSSP5
CC (AAB11710-B11711), and to nucleic acids encoding them (AAA61733-A61734).
CC The invention also relates to vectors and transformatants comprising BSSP5
CC nucleic acids; transgenic animals in which the expression level of BSSP5
CC can be varied; and an mBSSP5 knockout mouse. The invention additionally
CC encompasses anti-BSSP5 antibodies and methods of production of such
CC antibodies, methods of BSSP5 detection using the antibodies, and the
CC use of BSSP5 proteins or fragments as diagnostic markers for certain
CC medical conditions, e.g., pancreatitis. A method for detecting
CC pancreatitis comprising measuring BSSP5 concentration in the blood or
CC urine, and a pancreatitis diagnostic agent containing an anti-BSSP5
CC antibody is also disclosed. Nucleotides encoding BSSP5 were initially
CC isolated in a human brain cDNA library using degenerate PCR primers
CC (AAA61744-A61745) based on conserved regions of serine proteases. The
CC BSSP5 serine proteases and nucleotides encoding them are useful in
CC detecting homologues, mutants and polymorphic variants in biological
CC samples (e.g., blood, urine, brain, prostate gland, placenta, testis,
CC pancreas and spleen) as diagnostic markers for conditions such as
CC Alzheimer's disease, epilepsy, cancer, inflammation, infertility,
CC pancreatitis and prostatic hypertrophy. Sequence AAB11710 represents
CC human BSSP5 (hBSSP5), and sequence AAB11711 represents murine BSSP5
CC (mBSSP5).
XX
XX Sequence 264 AA;
XX
XX Query Match 89.5%; Score 1093; DB 21; Length 264;
XX Best Local Similarity 85.7%; Pred. No. 1.6e-87;
XX Matches 198; Conservative 21; Mismatches 12; Indels 0; Gaps 0;
QY 1 MLLSLTSLVLGSSWGGCIPAIKPAKPSQRIVNGENAVLGSWPQVSLQDSSGFHFC 60
DB 1 MLLSLTSLVLGSSWGGCIPAITPALTSTNQIRIVNGENAVPGSWPQVSLQDTEGFHFC 60
QY 61 GGSLSISQSWVVTAAHCNVSPGRHFVVLGEYDRSSNAEPQLQVLSVRAITHPSWNSTMMN 120
DB 61 GGSLSISPNWVVTAAHCNQVTPGRHFVVLGEYDRSSNAEPQVLSIARAITHPNNANTMMN 120
QY 121 DVTLLKLASPAQYTTTRISPVCLASSNEALTEGLTCVTTGWGRLSGVGNVTPAHLQOVALP 180
DB 121 DVTLLKLASPAQYTTTRISPVCLASSNEALTEGLTCVTTGWGRISGVGNVTPARLQOVALP 180
QY 181 LVTVNOCROYWDSSTITDSMICAGGAGASSCGDGGPLVCQKGTWVLIGI 231
DB 181 LVTVNOCROYWAGIITDAMICAGSGASSCGDGGPLVCQKGTWVLIGI 231
RESULT 3
AAB54077
ID AAB54077 standard; Protein; 192 AA.
XX
XX AAB54077;
XX
XX 09-MAR-2001 (first entry)
XX
XX Human pancreatic cancer antigen protein sequence SEQ ID NO:529.
XX

XX Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
KW detection; diagnosis; identification; cytostatic; neuroprotective;
KW neoplastic; immunomodulatory; relaxant; contraceptive; gynaecological;
KW antiinflammatory; cardiant; gene therapy; chromosome mapping;
KW linkage analysis; tissue identification; tissue typing; forensic;
KW neural; immune system; muscular; reproductive; gastrointestinal;
KW pulmonary; cardiovascular; renal; proliferative.
XX
OS Homo sapiens.
XX WO200055320-A1.
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US05989.
XX
XX 12-MAR-1999; 99US-0124270.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX WPI: 2000-579444/54.
XX N-PSDB; AAC98842.
XX
XX New nucleic acid that is a pancreatic cancer antigen for preventing,
PT treating, or ameliorating a medical condition, particular pancreatic
PT cancer, or for use in assays for diagnosing a pathological condition -
XX
XX Claim 11; Page 966; 1379pp; English.
XX
XX AAC98773 to AAC99231 encode the human pancreatic cancer associated
CC proteins, called pancreatic cancer antigens, given in AAB54008 to
CC AAB54466. The human pancreatic cancer antigens have cytostatic,
CC neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive,
CC gynaecological, cardiant and antiinflammatory activities, and can be used
CC in gene therapy. The polynucleotide and proteins can be used for
CC preventing, treating, or ameliorating a medical condition or in assays
CC for diagnosing a pathological condition or a susceptibility to one in a
CC subject. Binding partners to the proteins and the activity of the
CC proteins can be identified. The pancreatic cancer antigens can be used to
CC detect, treat or prevent pancreatic disorders, especially cancer.
CC Agonists and antagonists to the antigens can be screened for. The
CC pancreatic cancer antigen polynucleotides can be used to design nucleic
CC acid hybridisation probes that can be used in chromosome mapping, linkage
CC analysis, tissue identification and/or typing and a variety of forensic
CC and diagnostic methods. The proteins can be used to generate antibodies
CC which are used to purify, detect and target the polypeptides, including
CC both in vivo and in vitro diagnostic and therapeutic methods. The
CC proteins can be used to treat or prevent neural, immune system, muscular,
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
CC sequences used in the exemplification of the present invention.
XX
XX Sequence 192 AA;
SQ
Query Match 81.6%; Score 996; DB 21; Length 192;
Best Local Similarity 99.5%; Pred. No. 3.3e-79;
Matches 189; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 SUTLSLVLLGSSWGGCIPAIKPAKPSQIRIVNGENAVLGSPQVSLQDSSGFHFCGSL 64
DB 1 SUTLSLVLLGSSWGGCIPAIKPAKPSQIRIVNGENAVLGSPQVSLQDSSGFHFCGSL 60
QY 65 ISQSVVTTAAHCNVSGRHFVVLGEGYDRSSNAEPLOVLSVRAITHPSWNSTTNNNDVTL 124
DB 61 ISQSVVTTAAHCNVSGRHFVVLGEGYDRSSNAEPLOVLSVRAITHPSWNSTTNNNDVTL 120
QY 125 LKASPAQVTTTRISPVCLASSNEALTEGLTCVTTGNGRLSGVGNVTPAHLQVVALPLTV 184
DB 121 LKASPAQVTTTRISPVCLASSNEALTEGLTCVTTGNGRLSGVGNVTPAHLQVVALPLTV 180

QY 185 NOCROYWDSS 194
DB 181 NOCROYWDSS 190
RESULT 4
AAU82738
ID AAU82738 standard; Protein; 263 AA.
XX AAU82738;
XX 23-APR-2002 (first entry)
XX Amino acid sequence of novel human protease #37.
XX Human; protease; cancer; immune-related disorder; cardiovascular disease;
KW neuronal-associated disease; metabolic disorder; inflammatory disorder;
KW nervous system disorder; sexual dysfunction; pain; mood disorder;
KW hypertension; psychotic disorder; neurological disorder; dyskinesia;
KW viral infection; human immunodeficiency virus; HIV; non-viral infection;
KW ocular disease; cytostatic; enzyme.
XX Homo sapiens.
XX WO200200860-A2.
XX 03-JAN-2002.
XX 26-JUN-2001; 2001WO-US20171.
XX 26-JUN-2000; 2000US-214047P.
XX (SUGE-) SUGEN INC.
XX Plowman G, Whyte D, Sudarsanam S, Manning G, Caenepeel S;
PI Charyczak G;
XX WPI: 2002-139913/18.
XX N-PSDB; ABK31780.
XX Nucleic acids encoding novel human proteases, useful for useful for
PT treating diseases and disorders such as cancers, immune-related
PT diseases and disorders, cardiovascular disease (e.g. restenosis) and
PT inflammatory disorders -
XX
XX Claim 6; Fig 2M; 313pp; English.
XX The present invention relates to the isolation of novel human
XX proteases, and the nucleic acids encoding them. The sequences of
XX the invention are useful for treating diseases and disorders such as
XX cancers (e.g. breast, colon, lung), immune-related diseases and disorders
XX (e.g. inflammatory diseases and asthma), cardiovascular diseases
XX (e.g. restenosis and coronary thrombosis), brain or neuronal-associated
XX diseases, metabolic disorders (e.g. diabetes, obesity), inflammatory
XX disorders (e.g. rheumatoid arthritis and psoriasis), central or
XX peripheral nervous system diseases, migraines, pain, sexual dysfunction,
XX mood disorders, attention disorders, cognition disorders, hypotension,
XX hypertension, psychotic disorders, neurological disorders
XX (e.g. Alzheimer's disease, Parkinson's disease) and dyskinesias.
XX The nucleic acids and polypeptides are also useful for treating viral
XX infections caused by human immunodeficiency virus (HIV), and non-viral
XX infections such as ocular disease (e.g. glaucoma) and macular
XX degeneration. AAU82702-AAU82760 represent the novel human proteases of
XX the invention.
XX Sequence 263 AA;
SQ
Query Match 56.1%; Score 684.5; DB 23; Length 263;
Best Local Similarity 55.0%; Pred. No. 8e-52;
Matches 127; Conservative 36; Mismatches 67; Indels 1; Gaps 1;
QY 1 MULLSLTSLVLLGSSWGGCIPAIKPAKPSQIRIVNGENAVLGSPQVSLQDSSGFHFC 60

Db 1 MASLWLLSCFSLVGAAGCGVPAIHPVLSGLSRIVNGEDAVPGSWPQVSLQDKTGPHFC 60
QY 61 GGSLSISQVWVYTAARCNVSPGRHVPVLGEYDRSSNAEPLOVLSVSRATHPSWNSTMMN 120
Db 61 GGSLSISQVWVYTAARCNVSPGRHVPVLGEYDRSSNAEPLOVLSVSRATHPSWNSTMMN 119
QY 121 DVTLLKIASPAQYTRISPVCLASSNEALTEGLTCVTGWLGRSGVGNVTPAHLLQOVALP 180
Db 120 DITLLKLTAPRFSQTSVAVCLPSADDDPAGTLCATGKTKYNANKTPDKLQQAALP 179
QY 181 LVTYNOCROYWDSITSMICAGAGASSCGQSDSGGLVLCOKGNTWVLGI 231
Db 180 LLSNAECKSMGRITDVMICAGAGVSSCMGSDGGLVLCOKGNTWVLGI 230
RESULT 5
AAY99596
ID AAY99596 standard; protein; 245 AA.
XX AAY99596;
XX 13-SEP-2000 (first entry)
XX Bovine chymotrypsinogen A.
XX Bovine; plasminogen activator; cardiant; thrombolytic;
KW heart attack; stroke; blood clotting disorder.
XX Bos taurus.
XX W0200032759-A1.
XX 08-JUN-2000.
XX 06-MAY-1999; 99WO-US09991.
XX 02-DEC-1998; 98US-0110588.
XX (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX Lin X, Zhang XC, Tang JUN;
XX WPI; 2000-422975/36.
XX Polypeptide with plasminogen activator activity useful as thrombolytic
PT agent for treating blood clot disorders e.g. heart attack, comprises 10
PT amino acid peptide fragment for recognition or activation of
PT plasminogen
XX Disclosure; Page 35-36; 4lpp; English.
XX The present sequence is bovine chymotrypsinogen. It was included in
CC a review of sequence homologies of several plasminogen activators.
CC Plasminogen is the principal serine protease zymogen in the
CC extracellular fluids of vertebrates. Its active form, plasmin, is
CC implicated in pericellular proteolysis associated with a wide range of
CC physiological and pathological processes. Plasminogen expression is
CC regulated by plasminogen activators which hydrolyse a peptide bond in
CC plasminogen to convert it to plasmin or form tight binding complexes
CC with plasminogen to spontaneously convert it to plasmin. The sequence
CC homology analysis has identified a six amino acid peptide involved in
CC plasminogen activation. This peptide is particularly useful when inserted
CC between amino acid residues 644 and 645 of full length human
CC plasminogen. Novel plasminogen activators have been made based upon the
CC plasminogen activation/recognition site of plasminogen binding proteins.
CC The polypeptides are useful in preparing thrombolytic agents for treating
CC blood clotting disorders such as heart attack.
XX Sequence 245 AA;

Query Match 53.8%; Score 657.5; DB 21; Length 245;
Best Local Similarity 54.9%; Pred. No. 1.7e-49;
Matches 117; Conservative 37; Mismatches 58; Indels 1; Gaps 1;

QY 19 CGIPAIPALSFQRIVNGENAVLGSWPQVSLQDSSGFHFCGSLISQSWVYTAHCNV 78
Db 1 CGVPAIQVLSGLSRIVNGEAVPGSWPQVSLQDKTGPHFCGSLINENWVYTAHCGV 60
QY 79 SPGRHFVVLGEYDRSSNAEPLOVLSVSRATHPSWNSTMMNVDVTLKLPASPAQYTRIS 138
Db 61 TTS-DVVVAGEFDQSGSSERIQKIAKVPKNSKYNLSLTINNDITLLKLTAAASFQTS 119
QY 139 PVLASSNEALTEGLTCVTGWLGRSGVGNVTPAHLLQOVALPLVTVNOCROYWDSITDS 198
Db 120 AVCLPSASDDFAAGTTCVTGWLGRTRYTNANTPRLQOASLPLLSNTNCKRYMGTKRDA 179
QY 199 MICAGAGASSCGQSDSGGLVLCOKGNTWVLGI 231
Db 180 MICAGAGVSSCMGSDGGLVLCOKGNTWVLGI 212
RESULT 6
AAB98504
ID AAB98504 standard; Protein; 231 AA.
XX AAB98504;
XX 03-AUG-2001 (first entry)
XX Human chymotrypsin serine protease catalytic domain.
XX Human; TADG-15; cytostatic; vaccine; ovarian tumour; cancer;
KW tumour antigen-derived gene 15; serine protease; chymotrypsin.
XX Homo sapiens.
XX W0200129056-A1.
XX 26-APR-2001.
XX 20-OCT-2000; 2000WO-US29095.
XX 20-OCT-1999; 99US-0421213.
XX (UYAR-) UNIV ARKANSAS.
XX O'Brien TJ, Tanimoto H;
XX WPI; 2001-381031/40.
XX Novel extracellular serine protease, termed tumor antigen-derived gene
PT 15 protein overexpressed in carcinomas and DNA encoding it, for
PT diagnosis, treatment, prevention of cancer, particularly breast,
PT ovarian cancer
XX Example 10; Fig 1; 130pp; English.
XX The present invention relates to human tumour antigen-derived gene 15
CC (TADG-15) protein and coding sequence (see AAH23601 and AAB98500).
CC TADG-15 is an extracellular serine protease. It was found that TADG-15 is
CC over-expressed in ovarian tumours. TADG-15 protein or its fragments of
CC 9-20 residues that lack TADG-15 protease activity are useful for
CC vaccinating an individual against TADG-15, having, suspected of having or
CC at risk of getting cancer. Furthermore, the TADG-15 gene can be used as a
CC diagnostic or therapeutic target in cancer. The present sequence was used
CC in a sequence homology alignment with the catalytic domain of TADG-15.
XX Sequence 231 AA;

Query Match 50.9%; Score 621.5; DB 22; Length 231;
Best Local Similarity 56.8%; Pred. No. 2.2e-46;
Matches 113; Conservative 31; Mismatches 54; Indels 1; Gaps 1;

QY 33 RYVNGENAVLGSWPQVSLQDSSGFHFCGSLISQSWVYTAHCNVSPGRHFVVLGEYDR 92
Db 1 RYVNGENAVLGSWPQVSLQDSSGFHFCGSLISQSWVYTAHCNVSPGRHFVVLGEYDR 59

Qy 93 SSNAPLQVLSVRAITHPSWNTTMMNDVTLKLSAPQVTRISPVCLASSNEALTEG 152
 Db 60 GSDEENIQVLAKVFKPKSILTVNDITLLKATAPFSQTVSAVCLPSADDFPAG 119
 Qy 153 LTCVTTGNGRLSGVGNVTPAHLLQVALPLVTVNCQRYWDSITDSMICAGAGASSCOG 212
 Db 120 TLCAITGNGKTKYNANKTPDKLQQAALPLLSNAECKSKWGRRTITDVMICAGASGVSCMG 179
 Qy 213 DSGGPLVCOKGNWTVLIGI 231
 Db 180 DSGGPLVCOKDGAMTLVGI 198

RESULT 7
 AAP61724
 ID AAP61724 standard; Protein: 269 AA.

XX AC AAP61724;
 XX DT 23-OCT-1991 (first entry)
 XX DE Porcine elastase II.
 XX KW Colibacillus; yeast.
 XX OS Sus scrofa.
 XX PN JP61192289-A.
 XX PD 26-AUG-1986.
 XX PF 22-FEB-1985; 85JP-0034050.
 XX PR 22-FEB-1985; 85JP-0034050.
 XX PA (KIRI) KIRIN BREWERY KK.
 XX WPI: 1986-262895/40.
 XX N-PSDB; AAN60919.
 PT Biological prepn. of pig elastase II - using DNA with base
 PT sequence coding for specified aminoacid sequence.
 PS Claim 1; Fig 1; 11pp; Japanese.
 CC The porcine elastase product may be efficiently expressed from a
 CC transformed host such as colibacillus or yeast, yielding the product
 CC more efficiently than spleen extraction.
 CC See also J61192288.
 XX SQ Sequence 269 AA;

Query Match 38.7%; Score 472; DB 7; Length 269;
 Best Local Similarity 45.6%; Pred. No. 3e-33;
 Matches 109; Conservative 32; Mismatches 84; Indels 14; Gaps 9;

Qy 2 LLSLTLSLVLLGSSWGGCIPAIKPAISFQRIYVNGENAVLGSWPQVSLQ--DSSG--FH 58
 Db 1 MIRALLLTSLVAG--ALSCGLPANLPQ---PRVVGEDARPNSWPQVSLQYDSSGQWRH 56
 Qy 59 FCGGSLISQSVWVTAHCNVSPGRHFVVLGEYDRSSNAEPQLQVLSVRAITHPSWNTTMM 118
 Db 57 TCGGTLVDQSWLTAHCNISSTRTYRVVLRHSLSTNEPGSLAVKVLVHQDWNNSOL 116
 Qy 119 N--NDVTLLKLSAPQYTRISPVCLASSNEALTEGLTCVTTGWGRLSGVGNVTPAHLLQ 176
 Db 177 VALPLVTYNOCRQ--YWDSSITDSMICAGAG--ASSCGDGGPLVCOKGN--TWVLIGI 231
 Db 176 GOLLVVDYATCSKPGWGSTVTKTNMICAGDGIISSCNGDSGGPLNCQAGQWQVHGI 234

RESULT 8
 AAP70758
 ID AAP70758 standard; Protein: 269 AA.
 XX AC AAP70758;
 XX DT 19-APR-1991 (first entry)
 XX DE Pig pancreas elastase-2.
 XX KW Lipoprotein metabolism; lipid.
 XX OS Sus scrofa.

XX FH Key Location/Qualifiers
 XX FT Misc-difference 2..16
 XX FT /note= "May be absent"
 XX PN JP62000276-A.
 XX PD 06-JAN-1987.
 XX PF 25-JUN-1985; 85JP-0138494.
 XX PR 25-JUN-1985; 85JP-0138494.
 XX PA (SANY) SANKYO KK.
 XX WPI: 1987-040875/06.
 XX N-PSDB; AAN71121.

PT Pancreas elastase prodn. used to improve lipoprotein metabolism -
 PT comprises isolating RNA coding elastase, synthesising single and
 PT double chain cDNA and introducing recombinant into host
 XX PS Disclosure; Page 480; 18pp; Japanese.
 XX CC Sequence encoding pancreas elastase may be used to transform an
 XX CC expression system. The product is useful in rejuvenating the
 XX CC elasticity of the arterial wall, and improving unusual serum lipid
 XX CC levels and lipoprotein metabolism.

SQ Sequence 269 AA;

Query Match 38.7%; Score 472; DB 8; Length 269;
 Best Local Similarity 45.6%; Pred. No. 3e-33;
 Matches 109; Conservative 32; Mismatches 84; Indels 14; Gaps 9;

Qy 2 LLSLTLSLVLLGSSWGGCIPAIKPAISFQRIYVNGENAVLGSWPQVSLQ--DSSG--FH 58
 Db 1 MIRALLLTSLVAG--ALSCGLPANLPQ---PRVVGEDARPNSWPQVSLQYDSSGQWRH 56
 Qy 59 FCGGSLISQSVWVTAHCNVSPGRHFVVLGEYDRSSNAEPQLQVLSVRAITHPSWNTTMM 118
 Db 57 TCGGTLVDQSWLTAHCNISSTRTYRVVLRHSLSTNEPGSLAVKVLVHQDWNNSOL 116
 Qy 119 N--NDVTLLKLSAPQYTRISPVCLASSNEALTEGLTCVTTGWGRLSGVGNVTPAHLLQ 176
 Db 177 VALPLVTYNOCRQ--YWDSSITDSMICAGAG--ASSCGDGGPLVCOKGN--TWVLIGI 231
 Db 176 GOLLVVDYATCSKPGWGSTVTKTNMICAGDGIISSCNGDSGGPLNCQAGQWQVHGI 234

RESULT 9
 AAR29621
 ID AAR29621 standard; Protein: 253 AA.

XX AC AAR29621;
 XX DT 30-APR-1993 (first entry)

```

XX DE Porcine pancreatic elastase III.
XX KW Pig; swine; bile acid; gall bladder; bile acid secretion promoters;
XX KW liver function improvers; ss.
XX OS Sus scrofa domestica.
XX PN JP04325090-A.
XX PD 13-NOV-1992.
XX PF 23-APR-1991; 91JP-0092069.
XX PR 23-APR-1991; 91JP-0092069.
XX PA (SANY ) SANKYO CO LTD.
XX WPI; 1992-428826/52.
DR N-PSDB; AAQ31724.
XX PT Pig pancreas elastase protein - used in bile acid secretion
PT promoters and liver function improvers
XX PS Claim 1; Page 8; 10pp; Japanese.
XX CC This sequence represents porcine pancreatic elastase III. It was
CC expressed in E. coli YA21 using the expression vector pELF001.
CC It may be used in bile acid secretion promoters and liver function
CC improvers.
XX SQ Sequence 253 AA;

Query Match 36.6%; Score 446.5; DB 13; Length 253;
Best Local Similarity 43.5%; Pred. No. 4.8e-31;
Matches 97; Conservative 33; Mismatches 78; Indels 15; Gaps 8;

QY 19 CGIPAIKPAKLSFSQIRIVNGENAVLGSMPQVSLQ-DSSGF--HFCGGLISQSWVYTAH 75
DB 1 CGRPSYNPA----SRVVGEDAVPYSWPQVSLQYEKNGVFQHTCGGLIAPDWVLTAGH 56
QY 76 CNVSPGRHFVVLGEYDRSSNAEPLQVLSVRA--ITHPSWNSTMN--NDVTLKLASPA 131
DB 57 C1SSSTYQVVLGEYDRSENEGFEQVPIINAGDLFVHPRWNSNCVSCGNDIALVLSRSA 116
QY 132 QYTRISPVCLASNEALTEGLTCVTGWRGLSGVGNVTPPAHLQVALPLVTYNQCRQ-- 189
DB 117 QLGDKVQLACLPPAGDILPNDTPCYISGWRGLYTNGLP-PDKLQALLPVVDYQHCKWD 175
QY 190 YWDSSTIDSMICAGGAGASSCGDGGPLVCOKGN-TWVLIGI 231
DB 176 WVGSTVKQTWVCAGGDIRSCGNGDSGGLNCPAAGSGWQVHGV 218

RESULT 10
AAP60062
ID AAP60062 standard; Protein; 269 AA.
XX AC AAP60062;
XX AC AAP60062;
XX DT 06-SEP-1991 (first entry)
XX DE Sequence of human pancreatic elastase IIA encoded on pH2E2.
XX KW Enzyme; serum lipoprotein metabolism.
XX OS Homo sapiens.
XX PN EP198645-A.
XX PD 22-OCT-1986.
XX PF 07-APR-1986; 86EP-0302557.

```

```

XX 23-OCT-1985; 85JP-0236686.
PR 05-APR-1985; 85JP-0072308.
PR 27-APR-1985; 85JP-0091986.
PR 26-JUL-1985; 85JP-0163964.
PR 02-DEC-1985; 85JP-0271128.
XX (SANY ) SANKYO KK.
XX Takiguchi Y, Tani T, Kawashima I, Eurukawa H, Ohmine T;
PI Ohsumi J;
XX WPI; 1986-280300/43.
DR N-PSDB; AAN60076.
XX PT New genetically engineered human pancreatic elastase - obt'd.
PT using hosts modified DNA coding for enzyme
XX PS Example; Page 14-15; 45pp; English.
XX CC The genetically engineered prod. can eliminate the dependency on
CC human pancreas supplies for the elastase, and avoids antibody
CC formation and possibility of anaphylaxis using porcine elastase.
XX SQ Sequence 269 AA;

Query Match 36.4%; Score 445; DB 7; Length 269;
Best Local Similarity 43.9%; Pred. No. 6.9e-31;
Matches 105; Conservative 28; Mismatches 92; Indels 14; Gaps 8;

QY 2 LLSLTSLVLLGSSWCGGIPAIKPAKLSFSQIRIVNGENAVLGSMPQVSLQDSSG---FH 58
DB 1 MIRTLLSTIVAG-ALSCGDPTYP---YVTVYVGGEARPNPMPQVSLQYSSNGKWH 56
QY 59 FCGGLISQSWVYTAACHNVSPGRHFVVLGEYDRSSNAEPLQVLSVRAITHPSWNSTTM 118
DB 57 TCGGLIANSWVLTAAHCISSTRTYRVGLGRHLYVAESGLAVSYKIVVHKDWSNQI 116
QY 119 N--NDVTLKLASPAQYTRISPVCLASNEALTEGLTCVTGWRGLSGVGNVTPPAHLQ 176
DB 117 SKGNDIALLKANPVSILTQIQLACLPPAGTILPNNYPCVVTGWRGLQTNAGV-PDVLQ 175
QY 177 VALPLVTYNQCRQ--YWDSSITDSMICAGGAG-SSCGDGGPLVCOKGN-TWVLIGI 231
DB 176 GRLLVVDYATCSSAWWSSVKTSMICAGDGVISSCGDGGPLNCQASDGRWQVHGI 234

RESULT 11
AAP61723
ID AAP61723 standard; Protein; 269 AA.
XX AC AAP61723;
XX DT 23-OCT-1991 (first entry)
XX DE Human elastase II.
XX KW Colibacillus; yeast.
XX OS Homo sapiens.
XX PN JP61192288-A.
XX PD 26-AUG-1986.
XX PF 22-FEB-1985; 85JP-0034049.
XX PR 22-FEB-1985; 85JP-0034049.
XX (KIRI ) KIRIN BREWERY KK.
XX WPI; 1986-262894/40.
DR N-PSDB; AAN60706.

```

XX New nucleic acid that is a pancreatic cancer antigen for preventing,
PT treating, or ameliorating a medical condition, particular pancreatic
PT cancer, or for use in assays for diagnosing a pathological condition
XX

```

XX (SANY ) SANKYO KK.
XX WPI: 1987-040875/06.
XX N-PSDB; AAM71123.
XX
XX Pancreas elastase prodn. used to improve lipoprotein metabolism -
XX PT comprises isolating RNA coding elastase, synthesising single and
XX PT double chain cDNA and introducing recombinant into host
XX PT
XX PS Disclosure; Page 482; 18pp; Japanese.
XX
XX Sequence encoding pancreas elastase may be used to transform an
XX CC expression system. The product is useful in rejuvenating the
XX CC elasticity of the arterial wall, and improving unusual serum lipid
XX CC levels and lipoprotein metabolism.
XX
XX Sequence 269 AA;
XX
Query Match 35.0%; Score 427; DB 8; Length 269;
Best Local Similarity 41.0%; Pred. No. 2.6e-29;
Matches 98; Conservative 36; Mismatches 91; Indels 14; Gaps 8;
QY 2 LLSLSLSLVLLGSSWGGGIPAIKPAISQRIYVNGENAVLGSWPQVSLQDSSG---FH 58
Db 1 MIRTLLSLTVAG-ALSCGVSTYAPDMS---RMLGGEARPNPWPQVSLQYSSNGQWYH 56
QY 59 FCGSLISOSWVYTAACHNVSPCRHFVGLGEYDRSSNAEPLQVLSVRAITHPSWNSTTM 118
Db 57 TCGGSLIANSWVLTAAHCISSSRIYRVMLGQHNLYVAESGLAVSVSKIIVHDKWNSQV 116
QY 119 N--NDVTLLKLASPAQYTTTRISPVCLASSNEALTEGLTCVTTGWRLSGVGNVTPAHQQ 176
Db 117 SKGNDIALLKLANPVSLTDKIQLACLPPAGTILPNNPVCYVTGWRGLQTNG-ALPDDLKQ 175
QY 177 VALPLVTNQCRO--YWDSSITDSMICAGAGA-SSCOGDSGGPLVCQKGN-TWVLIGI 231
Db 176 GRLLVVDYATCSSSGWGWSTVKTNMICAGDGVICTCNGDSGGPLNCQASDGRWEVHGI 234

RESULT 14
AAM78338
ID AAM78338 standard; Protein; 1052 AA.
XX
AC AAM78338;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 1000.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
XX
PR 27-APR-2000; 2000US-0560875.
XX
PR 19-JUL-2000; 2000US-0598075.
XX
PR 01-SEP-2000; 2000US-0620325.
XX
PR 15-SEP-2000; 2000US-0654936.
XX
PR 20-OCT-2000; 2000US-0663561.
XX
PR 30-NOV-2000; 2000US-0693325.
XX
PA (HYPSE-) HYSEQ INC.

```

```

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
XX PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
XX PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX WPI: 2001-476283/51.
XX N-PSDB; AAK51471.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities,
XX useful in diagnosis and gene therapy -
XX
XX Claim 20; Page 3226-3228; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation.
XX
XX Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
XX (AAM80020) are omitted as the relevant pages from the sequence listing
XX were missing at the time of publication.
XX
XX Sequence 1052 AA;
XX
Query Match 35.0%; Score 427; DB 22; Length 1052;
Best Local Similarity 41.0%; Pred. No. 1.4e-28;
Matches 98; Conservative 36; Mismatches 91; Indels 14; Gaps 8;
QY 2 LLSLSLSLVLLGSSWGGGIPAIKPAISQRIYVNGENAVLGSWPQVSLQDSSG---FH 58
Db 1 MIRTLLSLTVAG-ALSCGVSTYAPDMS---RMLGGEARPNPWPQVSLQYSSNGQWYH 56
QY 59 FCGSLISOSWVYTAACHNVSPCRHFVGLGEYDRSSNAEPLQVLSVRAITHPSWNSTTM 118
Db 57 TCGGSLIANSWVLTAAHCISSSRIYRVMLGQHNLYVAESGLAVSVSKIIVHDKWNSQV 116
QY 119 N--NDVTLLKLASPAQYTTTRISPVCLASSNEALTEGLTCVTTGWRLSGVGNVTPAHQQ 176
Db 117 SKGNDIALLKLANPVSLTDKIQLACLPPAGTILPNNPVCYVTGWRGLQTNG-ALPDDLKQ 175
QY 177 VALPLVTNQCRO--YWDSSITDSMICAGAGA-SSCOGDSGGPLVCQKGN-TWVLIGI 231
Db 176 GRLLVVDYATCSSSGWGWSTVKTNMICAGDGVICTCNGDSGGPLNCQASDGRWEVHGI 234

RESULT 15
AAB36481
ID AAB36481 standard; Protein; 319 AA.
XX
AC AAB36481;
XX
DT 02-MAR-2001 (first entry)
XX
DE Fusion gene with human serine protease catalytic domain protein #12.
XX
KW Activation construct; catalytic; fusion gene; expression vector;
KW proteolysis; serine protease; zymogen precursor; characterisation;
KW analysis; modulator; identification.
XX
OS Homo sapiens.
XX
OS Synthetic.
XX
PN WO200066709-A2.
XX
PD 09-NOV-2000.
XX
PF 13-APR-2000; 2000WO-US09973.

```

XX 30-APR-1999: 99US-0303162.
XX (ORTH) ORTHO-MCNEIL PHARM RES INC.
XX PA
XX PI Darrow A, Qi J, Andrade-Gordon P;
XX DR WPI; 2000-687533/67.
XX
XX Expression vector for producing recombinantly producing serine protease
PT domains, comprising a prosequence, a prosequence, and a cloning site
PT for the insertion of catalytic domain cassette -
XX
XX
PS Disclosure; Page 47-50; 89pp; English.
XX
XX The present invention describes an expression vector (I) comprising in
CC frame and in order, a prosequence, a prosequence, and a cloning site
CC for the in frame insertion of catalytic domain cassette. (I) can be
CC used as a modulator of proteins expressed from a zymogen activation
CC construct. The recombinant catalytic domain of serine protease is
CC useful for identifying compounds modulating the activity of proteases
CC is expressed and activated from the zymogen activation construct. A
CC method from the present invention comprises combining a modulator of
CC the recombinant catalytic domain of a protease and measuring an effect
CC of the modulator on the protein preferably inhibiting or enhancing its
CC enzymatic activity or stimulation or inhibition of proteolysis mediated
CC by the expressed catalytic domain. The present sequence represents a
CC fusion gene with a human serine protease catalytic domain protein
CC sequence from the present invention.
XX
XX Sequence 319 AA;
Query Match 34.7%; Score 424; DB 21; Length 319;
Best Local Similarity 35.9%; Pred. No. 5.8e-29;
Matches 92; Conservative 39; Mismatches 97; Indels 28; Gaps 7;
QY 1 MLLSLTSLVLGSSWCGGIP-----AIKPALSFQRIVNGENAVLGSMWQVSL 51
Db 1 MAFLLWLLSCWALLGTTFGCGVPDYKDDDDAAALAAPFDDDDKIVGVYALEAGQWQVSI 60
QY 52 QDSGPHFCGSLISQSWVYTAHCNVSPGRHF-----VVLGEYDRSSNAEPLOVLSVR 106
Db 61 -TYEGVHVCGLSVSEQWVLSAAHC--FPSEHHKEAYEVKLGALHQLDYSYEDAKVSTLKD 117
QY 107 AITHPSNNTMNDVTLLKLSAPQVYTRISPVCLASSNEALTEGLTCVTTGWGRLS-G 165
Db 118 IIPHSYLVQESQGDIALQLSRPITFSRYIRPICLPAAANSPNGLHCTVTGWHVAPS 177
QY 166 VGNVTPAHLOQVALPLVTVNOCROYWDSS-----ITDSMICAG--GAGASSCQGDGSG 215
Db 178 VSLITPKPLQQLVETCLNIDAKPEEPHFVQDMVCAGYVEGGKDACQGDGSG 237
QY 216 GPLVCQKGNWTWVLGI 231
Db 238 GPLSCPVEGLWYLTGI 253

Search completed: December 20, 2002, 15:16:39
Job time : 37.6087 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:14:28 ; Search time 2.08696 Seconds
(without alignments)
1520.126 Million cell updates/sec

Title: US-09-856-319B-4_COPY_1_33

Perfect score: 166

Sequence: 1 MLLSLTSLVLLGSSWCGVPAITPALSYNQR 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	150	90.4	264	2 I38136	chymotrypsin-like
2	76	45.8	263	2 A21195	chymotrypsin (EC 3
3	73	44.0	263	2 A31299	chymotrypsin (EC 3
4	67	40.4	263	1 KYRTB	chymotrypsin (EC 3
5	55.5	33.4	835	2 F90260	hypothetical prote
6	55	33.1	231	2 A49265	flt3/flk-2 ligand
7	55	33.1	242	2 AE0906	monofunctional bio
8	55	33.1	268	2 S68826	pancreatic elastas
9	55	33.1	268	2 S68825	pancreatic elastas
10	55	33.1	466	2 T05664	hypothetical prote
11	52	31.3	219	2 B71511	probable sugar nuc
12	52	31.3	241	2 A10431	monofunctional bio
13	52	31.3	292	2 T10106	chitinase (EC 3.2.
14	52	31.3	557	2 JE0346	high-affinity carn
15	52	31.3	579	2 T04378	peptide transport
16	51	30.7	826	2 T02268	potassium transpor
17	51	30.7	1186	2 T03180	tyrosine protein k
18	50.5	30.4	384	2 AH0241	ceramide glucosylt
19	50	30.1	247	2 S23504	chymase (EC 3.4.21
20	50	30.1	260	2 S26043	chymase (EC 3.4.21
21	50	30.1	363	1 CBUTB	ubiquinol-cytochro
22	50	30.1	742	2 S55098	probable membrane
23	49.5	29.8	312	1 Q0E07	US10 protein - hum
24	49.5	29.8	484	2 T10255	nitrite transport
25	49.5	29.8	739	1 A70204	endopeptidase Clp
26	49.5	29.8	918	2 F86709	cation-transportin
27	49	29.5	94	2 A83108	hypothetical prote
28	49	29.5	104	2 A22706	vitelline membrane
29	49	29.5	141	2 F72121	ATP synthase, chai

ALIGNMENTS

RESULT 1

I38136

chymotrypsin-like proteinase (EC 3.4.21.-) CTRL-1 - human

C:Species: Homo sapiens (man)

C:Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 22-Jun-1999

C:Accession: I38136

R:Larsen, F.; Solheim, J.; Kristensen, T.; Kolsto, A.B.; Prydz, H.

Hum. Mol. Genet. 2, 1589-1595, 1993

A:Title: A tight cluster of five unrelated human genes on chromosome 16q22.1.

A:Reference number: I38135; MUID:94093544; PMID:8268911

A:Accession: I38136

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-264 <RES>

A:Cross-references: EMBL:X71874; NID:9406226; PIDN:CAA50710.1; PID:9406228

C:Genetics:

A:Gene: GDB:CTRL

A:Cross-references: GDB:204061

A:Map position: 16q22.1-16q22.1

A:Introns: 18/1; 52/3; 79/2; 106/3; 167/1; 211/3

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; serine proteinase

F:34-257/Domain: trypsin homology <TRY>

F:75,121,214/Active site: His, Asp, Ser #status predicted

Query Match 90.4%; Score 150; DB 2; Length 264;

Best Local Similarity 87.9%; Pred. No. 9.2e-12;

Matches 29; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLLSLTSLVLLGSSWCGVPAITPALSYNQR 33

Db 1 MLLSLTSLVLLGSSWCGGIPAIKPAISFSQR 33

RESULT 2

A21195

chymotrypsin (EC 3.4.21.1) 2 precursor - dog

C:Species: Canis lupus familiaris (dog)

C:Date: 27-Jul-1990 #sequence_revision 27-Jul-1990 #text_change 22-Jun-1999

C:Accession: A21195

R:Pinsky, S.D.; LaForge, K.S.; Luc, V.; Scheele, G.

Proc. Natl. Acad. Sci. U.S.A. 80, 7486-7490, 1983

A:Title: Identification of cDNA clones encoding secretory isoenzyme forms: sequence d

A:Reference number: A21195; MUID:84170253; PMID:6584866

A:Accession: A21195

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-263 <PIN>

A:Cross-references: GB:K01173; NID:9163945; PIDN:AAA30841.1; PID:9163946

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; protein digestion; serine proteinase

F:34-256/Domain: trypsin homology <TRY>

Query Match 40.4%; Score 67; DB 1; Length 263;
Best Local Similarity 44.8%; Pred. No. 0.34;
Matches 13: Conservative 5; Mismatches 11; Indels

A;Molecule type: mRNA

Matches 11; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis
A;Reference number: A71570; MUID:99000809; PMID:9784136
A;Accession: B71511
A;Status: preliminary
A:Molecule type: DNA
A:Residues: 1-219 <ARN>
A;Cross-references: GB:AE001320; GB:AE001273; NID:g3328891; PIDN:AA68062.1; PID:g3328891
A;Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: yd6P
C;Superfamily: sugar nucleotide phosphorylase

Query Match 31.3%; Score 52; DB 2; Length 219;
Best Local Similarity 36.4%; Pred. No. 23;
Matches 16; Conservative 3; Mismatches 11; Indels 14; Gaps 1;

QY 4 LSLTSLVLLGSSWG-----CGVPAITPALSYNQR 33
::: ||||| | || | ||
Db 1 MNLSCSLVLLGGKGRFNSLPQKQYTHLCGCEPLILHALHAYOR 44

RESULT 12
AI0431
monofunctional biosynthetic peptidoglycan transglycosylase (EC 2.4.2.-) [imported] - Yersinia pestis
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
C:Accession: AI0431
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-farraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AI0431
A;Status: preliminary
A:Molecule type: DNA
A:Residues: 1-241 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC92781.1; PID:g15981474; GSPDB:GN00175
C:Genetics:
A:Gene: mtgA
C;Keywords: glycosyltransferase; pentosyltransferase

Query Match 31.3%; Score 52; DB 2; Length 241;
Best Local Similarity 64.7%; Pred. No. 25;
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 17 WCGCVPAITPALSYNQR 33
|| | || |||||
Db 91 WGFVDGAIESALSHNQR 107

RESULT 13
TI0106
chitinase (EC 3.2.1.14) (class II, acidic) precursor - sweet orange
C:Species: Citrus sinensis (sweet orange)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: TI0106
R:Naïr, C.J.; Niedz, R.P.; Hearn, C.J.; Osswald, W.F.; Mayer, R.T. Biochim. Biophys. Acta 1351, 22-26, 1997
A;Title: cDNA cloning and expression of a class II acidic chitinase from sweet orange.
A:Reference number: Z16952; MUID:97236427; PMID:9116036
A;Accession: TI0106
A;Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-292 <NAI>
A;Cross-references: EMBL:Z70032; NID:g1220143; PIDN:CAA93847.1; PID:gi1220144
A;Experimental source: cv. Valencia; callus
C:Function:
A;Description: catalyzes the hydrolysis of the beta-1,4-linkages of N-acetyl-D-glucosamine
C;Superfamily: lectin-related plant chitinase; hevein chitin-binding domain homology; pectinase
C;Keywords: glycosidase; hydrolase; polysaccharide degradation
F;1-10/Domain: signal sequence #status predicted <SIG>
F;22-43/Domain: propeptide #status predicted <PRO>

Search completed: December 20, 2002, 15:19:49
Job time : 3.08696 secs

Search completed: December 20, 2002, 15:19:49
Job time : 3.08696 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:12:58 ; Search time 9.04348 seconds
(without alignments)
1210.790 Million cell updates/sec

Title: US-09-856-319B-4

Perfect score: 1418

Sequence: 1 MLLSLTLTLVLGSSWGG.....AMYTRVSKFSTWVQWYAYN 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1269	89.5	264	1 CTR2_HUMAN	P40313 homo sapien
2	832.5	58.7	263	1 CTR2_CANFA	P04813 canis faml
3	822.5	58.0	263	1 CTRB_HUMAN	P17596 homo sapien
4	798.5	56.3	263	1 CTRA_GADMO	P47796 gadus morhu
5	786.5	55.5	245	1 CTRB_BOVIN	P00767 bos taurus
6	781.5	55.1	263	1 CTRB_RAT	P07338 rattus norv
7	774.5	54.6	245	1 CTRA_BOVIN	P00766 bos taurus
8	748.5	52.8	245	1 CTRB_GADMO	P08646 gadus morhu
9	537	37.9	269	1 EL2_PIG	P08419 sus scrofa
10	520	36.7	269	1 EL2_BOVIN	P09461 bos taurus
11	517	36.5	271	1 EL2_RAT	P00774 rattus norv
12	515	36.3	269	1 EL2A_HUMAN	P08217 homo sapien
13	513.5	36.2	270	1 EL3B_HUMAN	P08861 homo sapien
14	503.5	35.5	253	1 CAC3_BOVIN	P05805 bos taurus
15	495.5	34.9	268	1 CLCR_HUMAN	P09895 homo sapien
16	492	34.7	271	1 EL2_MOUSE	P05208 mus musculu
17	486.5	34.3	268	1 CLCR_RAT	P55091 rattus norv
18	485.5	34.2	282	1 PLMN_MOUSE	P09018 mus musculu
19	484.5	34.2	810	1 EL3A_HUMAN	P09093 homo sapien
20	484	34.1	810	1 PLMN_HUMAN	P00747 homo sapien
21	479	33.8	269	1 EL2B_HUMAN	P08218 homo sapien
22	477	33.6	343	1 PLMN_SHEEP	P81286 ovis aries
23	475	33.5	810	1 PLMN_MACHU	P12545 macaca mula
24	474	33.4	790	1 PLMN_PIG	P06867 sus scrofa
25	470	33.1	333	1 PLMN_CANFA	P08009 canis faml
26	468.5	33.0	342	1 PSS8_RAT	P08087 rattus norv
27	468	33.0	454	1 TMS3_HUMAN	P57727 homo sapien
28	465	32.8	270	1 TRY1_MERIN	P50342 meriones un
29	463.5	32.7	273	1 MCT7_MOUSE	P02844 mus musculu
30	462.5	32.6	342	1 PSS8_MOUSE	P095d1 mus musculu
31	462.5	32.6	343	1 PSS8_HUMAN	P16651 homo sapien
32	460.5	32.5	275	1 TRY1_PIG	P092d1 sus scrofa
33	459	32.4	338	1 PLMN_HORSE	P80010 equus cabal

34	454.5	32.1	246	1 TRY1_RAT	P00762 rattus norv
35	452.5	31.9	273	1 MCT7_RAT	P27435 rattus norv
36	451.5	31.8	490	1 TMS2_MOUSE	Q91q8 mus musculu
37	447	31.5	231	1 TRYP_PIG	P00761 sus scrofa
38	447	31.5	266	1 EL1_BOVIN	Q28153 bos taurus
39	445.5	31.4	274	1 MCT6_RAT	P50343 bos taurus
40	444	31.3	266	1 EL1_RAT	P00773 rattus norv
41	443.5	31.3	275	1 TRY1_CANFA	P35944 canis faml
42	443	31.2	810	1 PLMN_ERIEU	Q29445 erinaceus e
43	442.5	31.2	246	1 TRY1_CANFA	P06871 canis faml
44	442	31.2	248	1 TRY1_CHICK	Q30627 gallus gall
45	441.5	31.1	812	1 PLMN_BOVIN	P06868 bos taurus

ALIGNMENTS

RESULT 1

ID	CTRL_HUMAN	STANDARD;	PRT;	264 AA.
AC	P40313;			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Chymotrypsin-like protease CTRL-1 precursor (EC 3.4.21.-).			
GN	CTRL OR CTRL1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=94093544; PubMed=8268911;			
RA	Larsen F., Solheim J., Kristensen T., Kolsto A.B., Prydz H.;			
RT	"A tight cluster of five unrelated human genes on chromosome			
RT	16q22.1.";			
RL	Hum. Mol. Genet. 2:1589-1595(1993).			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X71874; CAA50710.1; -			
DR	EMBL; X71877; CAA50711.1; -			
DR	HSSP; P00763; LDPO.			
DR	MEROPS; S01.256; -			
DR	Genew; HGNC:2524; CTRL.			
DR	MIN; I18888; -			
DR	InterPro; IPR001314; Chymotrypsin.			
DR	InterPro; IPR001254; Ser-protease_Try.			
DR	Pfam; PF00089; trypsin; 1			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	SMART; SM0020; TRYP-SPC; 1.			
DR	PROSITE; PS50240; TRYPIN_DOM; 1.			
DR	PROSITE; PS00134; TRYPIN_HIS; 1.			
DR	PROSITE; PS00135; TRYPIN_SER; 1.			
KW	Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal.			
FT	SIGNAL 1 18			POTENTIAL.
FT	PROPEP 19 33			ACTIVATION PEPTIDE (POTENTIAL).
FT	CHAIN 34 264			CHYMOTRYPSIN-LIKE PROTEASE CTRL-1.
FT	ACT_SITE 75 75			CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 121 121			CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 214 214			CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	CARBOHYD 114 114			N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID 19 141			BY SIMILARITY.
FT	DISULFID 60 76			BY SIMILARITY.
FT	DISULFID 155 220			BY SIMILARITY.
FT	DISULFID 187 201			BY SIMILARITY.

FT DISULFID 210 239 BY SIMILARITY.
SQ SEQUENCE 264 AA; 28002 MW; 3FG29F02FA6DDFB4 CRC64;

Query Match 89.5%; Score 1269; DB 1; Length 264;
Best Local Similarity 86.4%; Pred. No. 6.6e-102;
Matches 228; Conservative 25; Mismatches 11; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGSGGCGVPAITPALSYNQRIVNGENAVPGSWPQVSLQDNTGFHFC 60
DB 1 MLLSLTSLVLGSGGCGVPAITPALSYNQRIVNGENAVPGSWPQVSLQDNTGFHFC 60

QY 61 GGSLSIPNVVTAACQVTPGRHFVYVGLGYDRSSNAEPVQLSIAIRAITHPNMNMN 120
DB 61 GGSLSISQSWVYTAACNVSPGRHFVYVGLGYDRSSNAEPVQLSIAIRAITHPNMNMN 120

QY 121 DLTLLKLPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQOVVLP 180
DB 121 DVTLLKLPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQOVVLP 180

QY 181 LVTNOCROYWGARTIDAMICAGSGGCGSGGGLVPCQKGTWVLLIGIYSWGFKNCN 240
DB 181 LVTNOCROYWGARTIDAMICAGSGGCGSGGGLVPCQKGTWVLLIGIYSWGFKNCN 240

QY 241 IQAPAMYTRVSKFSTWVNOVAYN 264
DB 241 VRAPAVYTRVSKFSTWVNOVAYN 264

RESULT 2
CTR2_CANFA
ID CTR2_CANFA STANDARD; PRT; 263 AA.
AC P04813;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsinogen 2 precursor (EC 3.4.21.1).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84170253; PubMed=6584866;
RA Pinsky S.D., Laforce K.S., Luc V., Scheele G.;
RT "Identification of cDNA clones encoding secretory isoenzyme forms:
RT sequence determination of canine pancreatic chymotrypsinogen 2
RT mRNA."
RL Proc. Natl. Acad. Sci. U.S.A. 80:7486-7490(1983).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
CC Phe-I-Xaa, Leu-I-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; K01173; AAA30841.1; -
DR PIR; A21195; A21195.
DR HSP; P00766; IACH.
DR MEROPS; S01.152; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS02040; TRYPsin DOM; 1.
DR PROSITE; PS00134; TRYPsin H1S; 1.
DR PROSITE; PS00135; TRYPsin SER; 1.

KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
FT SIGNAL 1 18
FT CHAIN 19 263 CHYMOTRYPSINOGEN 2.
FT CHAIN 19 31 CHYMOTRYPSIN 2, A CHAIN.
FT CHAIN 34 164 CHYMOTRYPSIN 2, B CHAIN.
FT CHAIN 157 263 CHYMOTRYPSIN 2, C CHAIN.
FT ACT_SITE 75 75 CHARGE RELAY SYSTEM.
FT ACT_SITE 120 120 CHARGE RELAY SYSTEM.
FT ACT_SITE 213 213 CHARGE RELAY SYSTEM.
FT DISULFID 19 140 BY SIMILARITY.
FT DISULFID 60 76 BY SIMILARITY.
FT DISULFID 154 219 BY SIMILARITY.
FT DISULFID 186 200 BY SIMILARITY.
FT DISULFID 209 238 BY SIMILARITY.
SQ SEQUENCE 263 AA; 27787 MW; 2A2F449D813B3961 CRC64;

Query Match 58.7%; Score 832.5; DB 1; Length 263;
Best Local Similarity 57.2%; Pred. No. 1.9e-64;
Matches 151; Conservative 38; Mismatches 74; Indels 1; Gaps 1;

QY 1 MLLSLTSLVLGSGGCGVPAITPALSYNQRIVNGENAVPGSWPQVSLQDNTGFHFC 60
DB 1 MAFLWLLSCFALLGTAAGCGVPAIQVLSGLSRIVNGEDAVPGSWPQVSLQDNTGFHFC 60

QY 61 GGSLSIPNVVTAACQVTPGRHFVYVGLGYDRSSNAEPVQLSIAIRAITHPNMNMN 120
DB 61 GGSLSIEDWVTAACGCVRT-THQVVAGEFDQGSDAESIQVLKIAKVFKNPKFMETINN 119

QY 121 DLTLLKLPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQOVVLP 180
DB 121 DLTLLKLPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQOVVLP 179

QY 181 LVTNOCROYWGARTIDAMICAGSGGCGSGGGLVPCQKGTWVLLIGIYSWGFKNCN 240
DB 181 LLSNAECKKFWGSKIITDLAVCAGASVSCMGDSGLVPCQKDCANTLVIGIYSWGSGTCS 239

QY 241 IQAPAMYTRVSKFSTWVNOVAYN 264
DB 240 TSTPGVYARVTKLIPWVQILQAN 263

RESULT 3
CTR2_HUMAN
ID CTR2_HUMAN STANDARD; PRT; 263 AA.
AC P17538;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsinogen B precursor (EC 3.4.21.1).
GN CTRB1 OR CTRB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Pancreas;
RX MEDLINE=89134264; PubMed=2917002;
RA Tomita N., Izumoto Y., Horii A., Doi S., Yokouchi H., Ogawa M.,
RA Mori T., Matsubara K.;
RT "Molecular cloning and nucleotide sequence of human pancreatic
RT chymotrypsinogen cDNA."
RL Biochem. Biophys. Res. Commun. 158:569-575(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
CC Phe-I-Xaa, Leu-I-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announcements/](http://www.isb-sib.ch/announcements) or send an email to license@isb-sib.ch).

EMBL; M24400; AAA52128.1; --
 EMBL; BC005385; AAA05385.1; --
 PIR; A31299; A31299.
 HSP; P00766; ICHG.
 MEROPS; S01.152; --
 Genew; HGNC:2521; CTRB1.
 MW; 118890; --
 InterPro; IPR001314; Chymotrypsin.
 InterPro; IPR001254; Ser_protease_Try.
 Pfam; PF00089; trypsin; 1.
 PRINTS; PR00722; CHYMOTRYPSIN.
 SMART; SM00020; Tryp_Spc; 1.
 PROSITE; PS02040; TRYPSIN_DOM; 1.
 PROSITE; PS00134; TRYPSIN_HIS; 1.
 PROSITE; PS00135; TRYPSIN_SER; 1.
 Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
 SIGNAL 1 18
 CHAIN 19 263 CHYMOTRYPSINOGEN B.
 CHAIN 19 31 CHYMOTRYPSIN B, A CHAIN.
 CHAIN 34 164 CHYMOTRYPSIN B, B CHAIN.
 CHAIN 167 263 CHYMOTRYPSIN B, C CHAIN.
 ACT_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).
 ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
 ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
 DISULFID 19 140 BY SIMILARITY.
 DISULFID 60 76 BY SIMILARITY.
 DISULFID 154 219 BY SIMILARITY.
 DISULFID 186 200 BY SIMILARITY.
 DISULFID 209 238 BY SIMILARITY.
 SEQUENCE 263 AA; 27870 MW; 4C1C055A490B8701 CRC64;

Query Match 58.08; Score 822.5; DB 1; Length 263;
 Best Local Similarity 57.28; Pred. No. 1.4e-63;
 Matches 151; Conservative 38; Mismatches 74; Indels 1; Gaps 1;
 1 MLLSLTSLVLLGSSWCGVPAITPALSYNORIVNGENAVPGSWPMQVSLQDNTGPHFC 60
 1 MAFWLWSCWALLGTTFCGVPALHPVLSGLSRVINGEDAVPGSWPMQVSLQDNTGPHFC 60
 61 GGSLISPNVWVTAARCOVTPGRHFVVLGEYDRSSNAEPVQVLSIARATHPNWNTMNN 120
 61 GGSLISPDVWVTAARCGVRTS-DVVVAGEFDGSDDEENIQVLKIAKVPKPKFSILTNN 119
 121 DTLKLLKASPARYAQAQSPVCLASTNEALPSGLTCVTGTGWRISGVGNVTPARLQOVVLP 180
 120 DITLLKATPARFSQTSVAVCLPSADDDFPAGTLCATTGWGKTRYNANKTPDKLQQAALP 179
 181 LVTYNOCROYAGARTDAMICAGSGSSCGDGGPLVLCOKGNTWVLIGIVSGTKNCN 240
 180 LLSNAECKSGWRRTDVMICAGASGVSSCMGDSGGPLVLCOKDGAWLVLGVSGSDFTCS 239
 241 IQAPAMYTRVSKFSTWVQVMAYN 264
 240 TSSPGVYARVTKLIPWQKILAN 263

RESULT 4
 CTRA_GADMO STANDARD; PRT; 263 AA.
 AC P47796;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Chymotrypsin A precursor (EC 3.4.21.1).
 OS Gadus morhua (Atlantic cod).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
 RN NCBI_TaxID=8049;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pyloric caeca;
 RX MEDLINE=94368860; PubMed=8086467;
 RA Gudmundsdottir A., Oskarsson S., Eakin A.E., Craik C.S.,
 RA Bjarnason J.B.;
 RT "Atlantic cod cDNA encoding a psychrophilic chymotrypsinogen.";
 RL Biochim. Biophys. Acta 1219:211-214(1994).
 CC [2]
 CC SEQUENCE OF 19-30 AND 34-49.
 CC TISSUE=Pyloric caeca;
 RX MEDLINE=92111252; PubMed=1764912;
 RA Ageirsson B., Bjarnason J.B.;
 RT "Structural and kinetic properties of chymotrypsin from Atlantic cod
 (Gadus morhua). Comparison with bovine chymotrypsin.";
 RL Comp. Biochem. Physiol. 99A:327-335(1991).
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
 CC Phe-I-Xaa, Leu-I-Xaa.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcements/>
 CC or send an email to license@isb-sib.ch).

EMBL; X78490; CAA55242.1; --
 HSP; P00766; ICHG.
 MEROPS; S01.152; --
 InterPro; IPR001314; Chymotrypsin.
 InterPro; IPR001254; Ser_protease_Try.
 Pfam; PF00089; trypsin; 1.
 PRINTS; PR00722; CHYMOTRYPSIN.
 SMART; SM00020; Tryp_Spc; 1.
 PROSITE; PS02040; TRYPSIN_DOM; 1.
 PROSITE; PS00134; TRYPSIN_HIS; 1.
 PROSITE; PS00135; TRYPSIN_SER; 1.
 Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
 SIGNAL 1 18
 CHAIN 19 263 CHYMOTRYPSIN A.
 ACT_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).
 ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
 ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
 DISULFID 19 140 BY SIMILARITY.
 DISULFID 60 76 BY SIMILARITY.
 DISULFID 154 219 BY SIMILARITY.
 DISULFID 186 200 BY SIMILARITY.
 DISULFID 209 238 BY SIMILARITY.
 CONFLICT 21 21 R -> S (IN REF. 2).
 CONFLICT 25 25 S -> Q (IN REF. 2).
 CONFLICT 29 29 T -> S (IN REF. 2).
 CONFLICT 44 44 S -> T (IN REF. 2).
 CONFLICT 46 46 S -> Y (IN REF. 2).
 SEQUENCE 263 AA; 28294 MW; 47AAC699A0A64FBB CRC64;
 Query Match 56.38; Score 798.5; DB 1; Length 263;
 Best Local Similarity 56.98; Pred. No. 1.6e-61;
 Matches 144; Conservative 38; Mismatches 70; Indels 1; Gaps 1;
 12 LLGSSWCGVPAITPALSYNORIVNGENAVPGSWPMQVSLQDNTGPHFCGSLISPNWV 71
 12 LFRRTYGGCRPAISVITGYSRIVNGEAEVPHSWQVSLQDQTFHFHCGSLINWV 71
 72 TAAHCQVTPGRHFVVLGEYDRSSNAEPVQVLSIARATHPNWNTMNDTLKLLASPA 131
 72 TAAHCNV-KNVRVVLGEHRRSSNSEGQVMTVGQVFKPRYNGFTINNDILLVKLATPA 130

QY 132 RYTAQVSPVCLASTNEALPSGLTCTVTGMRISGVGNVTPARLQVVLPLVTNOCROYW 191
 Db 131 TLNMRVSPVCLASTDDVFEKKKCVTSGLTFYNNADTFALLOALPLLTNEQCKKFW 190
 QY 192 GARITDAMICAGGASSCGGSGPLVCGKNTWYLGIVSGTKNCNIQAPAMVTRYS 251
 Db 191 GNKISLMTICAGAAGASCGGSGPLVCGKAGSWLGVISWGSCTCTPMPGVYARVT 250
 QY 252 KFTSWINQVWAYN 264
 Db 251 ELRAWVDQITAAAN 263

RESULT 5
 CTRB_BOVIN STANDARD; PRT; 245 AA.
 AC P0767;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DE Chymotrypsinogen B (EC 3.4.21.1).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE, DISULFIDE BONDS, AND ACTIVE SITE.
 RX MEDLINE=68238908; PubMed=5649671;
 RA "Structure of chymotrypsinogen B compared with chymotrypsinogen A and
 trypsinogen.";
 RL Nature 218:343-346(1968).
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
 Phe-I-Xaa, Leu-I-Xaa.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- DATABASE: NAME=worthington enzyme manual;
 CC WWW="http://www.worthington-biochem.com/manual/C/CHY.html".
 DR PIR: A00953; KYBOB.
 DR HSSP: P00766; IACB.
 DR MEROPS: S01.152; -.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; TRYD_SPC; 1.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen.
 FT CHAIN 1 13 CHYMOTRYPSIN B, A CHAIN.
 FT CHAIN 15 146 CHYMOTRYPSIN B, B CHAIN.
 FT CHAIN 149 245 CHYMOTRYPSIN B, C CHAIN.
 FT ACT_SITE 57 57 CHARGE RELAY SYSTEM.
 FT ACT_SITE 102 102 CHARGE RELAY SYSTEM.
 FT ACT_SITE 195 195 CHARGE RELAY SYSTEM.
 FT DISULFID 1 122
 FT DISULFID 42 58
 FT DISULFID 136 201
 FT DISULFID 168 182
 FT DISULFID 191 220
 SQ SEQUENCE 245 AA; 25755 MW; 678016446FF5FEB5 CRC64;

Query Match 55.5%; Score 786.5; DB 1; Length 245;
 Best Local Similarity 57.7%; Pred. No. 1.6e-60;
 Matches 142; Conservative 32; Mismatches 71; Indels 1; Gaps 1;
 QY 19 CGVPALTALPSNORVINGENAVPGSWPQVSLQDNTGFHFCGSGSLISPNWVYTAACOV 78
 Db 1 CGVPALTQVLSGLARVINGEDAVPGSWPQVSLQDNTGFHFCGSGSLISDWDVYTAACOV 60

QY 79 TPGHEFVVLGEYDRSSNABPQVLSIARATTHPNMNMANTMNDTLKLPASARYTAQVS 138
 Db 61 TTS-DVVAGEEDQLEDTQVLKIGKVKFKPFSILVTRNDITLKLATPAQSETVS 119
 QY 139 PVLASTNEALPSGLTCTVTGMRISGVGNVTPARLQVVLPLVTNOCROYWGARITDA 198
 Db 120 AVCLPSADEDFPAGMLCATGKTKYNALKTDPKLLQATLPVSNTRDKRKYGSRVTDV 179
 QY 199 MICAGSGASSCGGSGPLVCGKNTWYLGIVSGTKNCNIQAPAMVTRYSKFTSWIN 258
 Db 180 MICAGSGVSCMGDSGGPLVCGKNGAWLAGIVSGSSTCTSTPFAVTRVATLMPWVQ 239
 QY 259 QVMAYN 264
 Db 240 ETIAAN 245
 RESULT 6
 CTRB_RAT STANDARD; PRT; 263 AA.
 AC P07338;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DE Chymotrypsinogen B precursor (EC 3.4.21.1).
 GN CTRB1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85054881; PubMed=6209274;
 RA Bell G.I., Quinto C., Quiroga M., Valenzuela P., Craik C.S.,
 Rutter W.J.;
 RT "Isolation and sequence of a rat chymotrypsin B gene.";
 RL J. Biol. Chem. 259:14265-14270(1984).
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
 Phe-I-Xaa, Leu-I-Xaa.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).
 DR EMBL: K02298; AAA98732.1; -.
 DR PIR: A22658; KYRTB.
 DR HSSP: P00766; 1CHG.
 DR MEROPS: S01.152; -.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; TRYD_SPC; 1.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 263 CHYMOTRYPSIN B.
 FT CHAIN 19 31 CHYMOTRYPSIN B, A CHAIN.
 FT CHAIN 34 164 CHYMOTRYPSIN B, B CHAIN.
 FT CHAIN 167 263 CHYMOTRYPSIN B, C CHAIN.
 FT ACT_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 19 140 BY SIMILARITY.
 FT DISULFID 60 76 BY SIMILARITY.


```
FT DISULFID 154 219 BY SIMILARITY.
FT DISULFID 186 200 BY SIMILARITY.
FT DISULFID 209 238 BY SIMILARITY.
SQ SEQUENCE 263 AA; 27849 MW; ACAFDACF8C4DA6D CRC64;

Query Match 55.1%; Score 781.5; DB 1; Length 263;
Best Local Similarity 53.8%; Pred. No. 4.6e-60;
Matches 142; Conservative 40; Mismatches 81; Indels 1; Gaps 1;

QY 1 MLLSLTSLVLLGSSWCGGCPAIPALSYNQRIYNGENAVPGSPWQVSLQDNTGFHFC 60
DB 1 MAFLWLVSCFALVGATGCGVPTIQVLTLGSLRIYNGDAIPGSPWQVSLQDNTGFHFC 60
QY 61 GGLISPNWVYTAHCQVTPGRHFVVLGELYDRSSNAEPVQVLSIARATHTPNWNAANTWNN 120
DB 61 GGLISEDWVYTAHCQVTKTS-DVVVAGEFFDGSDEEMIQVLKIAQVFNKFNFTVYRN 119
QY 121 DITLLKASPARYTAQVSPVCLASTNEALPSGLTCVTTGWGRISGVNVTFARLQOVVLP 180
DB 121 DITLLKATPAQFSETVSACVLPNVDDPPPGTVCAITGWGKTKYNALKTPEKIQQAALP 179
QY 181 LVTYNQCRQYWCARITDAMICAGSGASSCGDGGPLVCQKQNTWVLIGIVSWGTKNKN 240
DB 181 IVSEADCKKSGSKITDYMTTCAGASGVSSCMGDSGGPLVCQKQDGYVTLAGIVSWGVC 239
QY 241 IQAPAMYTRVSKFSTWVQVWVYNN 264
DB 240 TSTPAVYSRVTAIMPVWQVQILEAN 263

RESULT 7
CTRA_BOVIN STANDARD; PRT; 245 AA.
ID AC P00766;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsinogen A (EC 3.4.21.1).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE, DISULFIDE BONDS, AND ACTIVE SITE.
RX MEDLINE=67181721; PubMed=5971783;
RA Brown J.R., Hartley B.S.;
RT "Location of disulphide bridges by diagonal paper electrophoresis.
RT The disulphide bridges of bovine chymotrypsinogen A.";
RL Biochem. J. 101:214-228(1966).
RN [2]
RP REVISION TO 102.
RX MEDLINE=69106266; PubMed=5764436;
RA Blow D.M., Birktoft J.J., Hartley B.S.;
RT "Role of a buried acid group in the mechanism of action of
RT chymotrypsin.";
RL Nature 221:337-340(1969).
RN [3]
RP PRELIMINARY SEQUENCE.
RA Hartley B.S.;
RT "Amino-acid sequence of bovine chymotrypsinogen-A.";
RL Nature 201:1284-1287(1964).
RN [4]
RP SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=67183948; PubMed=5972866;
RA Meloun B., Klueh I., Kostka V., Moravek L., Prusik Z., Vanacek J.,
RA Kell B., Sorm F.;
RT "Covalent structure of bovine chymotrypsinogen A.";
RL Biochim. Biophys. Acta 130:543-546(1966).
RN [5]
RP ACTIVE SITE.
RX MEDLINE=67181723; PubMed=5971785;
RA Smillie L.B., Hartley B.S.;
```

```
RT "Histidine sequences in the active centres of some 'serine'
RT proteinases.";
RL Biochem. J. 101:232-241(1966).
RN [6]
RP X-RAY CRYSTALLOGRAPHY.
RX MEDLINE=72035052; PubMed=4399050;
RA Birktoft J.J., Blow D.M., Henderson R., Steitz T.A.;
RT "I. Serine proteinases. The structure of alpha-chymotrypsin.";
RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 257:67-76(1970).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF CHYMOTRYPSINOGEN.
RX MEDLINE=70177557; PubMed=5442169;
RA Freer S.T., Kraut J., Robertus J.D., Wright H.T., Xuong N.H.;
RT "Chymotrypsinogen: 2.5-A crystal structure, comparison with alpha-
RT chymotrypsin, and implications for zymogen activation.";
RL Biochemistry 9:1997-2009(1970).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF GAMMA-CHYMOTRYPSIN.
RX MEDLINE=82078042; PubMed=6914398;
RA Cohen G.H., Silverton E.W., Davies D.R.;
RT "Refined crystal structure of gamma-chymotrypsin at 1.9-A resolution.
RT Comparison with other pancreatic serine proteases.";
RL J. Mol. Biol. 148:449-479(1981).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (1.68 ANGSTROMS) OF ALPHA-CHYMOTRYPSIN.
RX MEDLINE=86011575; PubMed=4046030;
RA Tsukada H., Blow D.M.;
RT "Structure of alpha-chymotrypsin refined at 1.68-A resolution.";
RL J. Mol. Biol. 184:703-711(1985).
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
CC Phe-I-Xaa, Leu-I-Xaa.
CC -|- SUBCELLULAR LOCATION: Extracellular.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -|- DATABASE: NAME=Worthington enzyme manual;
CC WWW="http://www.worthington-biochem.com/manual/C/CHY.html".
DR PIR; A00952; KYBOA.
DR PDB; 2CGA; 15-APR-90.
DR PDB; 2CHA; 31-MAY-84.
DR PDB; 4CHA; 29-OCT-85.
DR PDB; 5CHA; 16-OCT-87.
DR PDB; 6CHA; 16-OCT-87.
DR PDB; 1CHG; 27-JAN-84.
DR PDB; 1CHO; 16-JUL-88.
DR PDB; 2GCH; 31-MAY-84.
DR PDB; 3GCH; 15-OCT-92.
DR PDB; 4GCH; 15-OCT-90.
DR PDB; 5GCH; 15-OCT-90.
DR PDB; 6GCH; 15-OCT-90.
DR PDB; 7GCH; 15-OCT-90.
DR PDB; 8GCH; 15-JUL-93.
DR PDB; 1GCT; 15-OCT-91.
DR PDB; 2GCT; 15-OCT-91.
DR PDB; 3GCT; 15-OCT-91.
DR PDB; 1ACB; 31-OCT-93.
DR PDB; 1GMC; 31-OCT-93.
DR PDB; 1GMD; 31-OCT-93.
DR PDB; 1CGI; 30-APR-94.
DR PDB; 1CGJ; 30-APR-94.
DR PDB; 1CGD; 22-JUN-94.
DR PDB; 1GHB; 22-JUN-94.
DR PDB; 1GHI; 22-JUN-94.
DR PDB; 1GMH; 30-SEP-94.
DR PDB; 2GMT; 01-NOV-94.
DR PDB; 1MTN; 17-AUG-96.
DR PDB; 1AB9; 20-AUG-97.
DR PDB; 1AFQ; 17-SEP-97.
DR PDB; 1CAO; 23-JUL-97.
DR PDB; 1CBW; 23-JUL-97.
DR PDB; 1VGC; 12-NOV-97.
DR PDB; 2VGC; 12-NOV-97.
DR PDB; 3VGC; 12-NOV-97.
DR PDB; 4VGC; 12-NOV-97.
DR PDB; 1HJA; 14-JAN-98.
```

```
DR MEROPS; S01.001; -
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser. protease_Try.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS00240; TRYP_SIN_DOM; 1.
DR PROSITE; PS00134; TRYP_SIN_HIS; 1.
DR PROSITE; PS00135; TRYP_SIN_SER; 1.
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen;
KW 3D-structure.
FT CHAIN 1 13 CHYMOTRYPSIN A, A CHAIN.
FT CHAIN 16 146 CHYMOTRYPSIN A, B CHAIN.
FT CHAIN 149 245 CHYMOTRYPSIN A, C CHAIN.
FT ACT_SITE 57 57 CHARGE RELAY SYSTEM.
FT ACT_SITE 102 102 CHARGE RELAY SYSTEM.
FT ACT_SITE 195 195 CHARGE RELAY SYSTEM.
FT DISULFID 1 122
FT DISULFID 42 58
FT DISULFID 136 201
FT DISULFID 168 182
FT DISULFID 191 220
FT HELIX 12 15
FT STRAND 20 21
FT TURN 24 25
FT TURN 28 29
FT STRAND 30 34
FT STRAND 40 46
FT STRAND 51 54
FT TURN 57 58
FT TURN 62 63
FT STRAND 65 68
FT TURN 69 69
FT STRAND 72 72
FT TURN 73 74
FT STRAND 81 90
FT TURN 92 93
FT STRAND 95 95
FT TURN 96 99
FT STRAND 100 100
FT TURN 101 101
FT STRAND 104 108
FT STRAND 122 122
FT TURN 126 127
FT TURN 132 133
FT STRAND 135 140
FT TURN 146 148
FT STRAND 154 154
FT STRAND 156 163
FT HELIX 165 168
FT TURN 169 172
FT HELIX 173 175
FT STRAND 180 184
FT TURN 192 193
FT TURN 195 196
FT STRAND 198 203
FT TURN 204 205
FT STRAND 206 216
FT TURN 218 219
FT TURN 222 223
FT STRAND 225 230
FT TURN 231 233
FT TURN 234 234
FT HELIX 235 243
SQ SEQUENCE 245 AA; 25666 MW; 91A9F28E2F3E3142 CRC64;

Query Match 54.6%; Score 774.5; DB 1; Length 245;
Best Local Similarity 57.3%; Pred. No. 1.7e-59;
Matches 141; Conservative 32; Mismatches 72; Indels 1; Gaps 1;

QY 19 CGVPAITPALSYNQIRVINGENAVPGSWPQVSLQDNTGPHFCGSGSLISPNWVYTAACQV 78
DB 1 CGVPAITPALSYNQIRVINGENAVPGSWPQVSLQDNTGPHFCGSGSLISPNWVYTAACQV 60
```



```

RC MEDLINE=98079203; PubMed=9418008;
RA Gestin M., le Huerou-Luron I., Wicker-Planquart C., le Drian G.,
RA Chai J.C., Puigserver A., Guilloteau P.
RT "Bovine pancreatic preproelastases I and II: comparison of nucleotide
RT and amino acid sequences and tissue specific expression.";
RC Comp. Biochem. Physiol. 118B:181-187(1997).
RL CL -!- FUNCTION: ACTS UPON ELASTIN.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Leu|-Xaa, Met|-Xaa
CC and Phe|-Xaa. Hydrolyzes elastin.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: PANCREAS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X97635; CAA66231.1; -.
DR HSP; P00766; ICHG.
DR MEROPS; S01.155; -.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS0240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolase; Serine protease; Zymogen; Signal.
FT SIGNAL 1 16
FT PROPEP 17 28 ACTIVATION PEPTIDE (BY SIMILARITY).
FT CHAIN 29 269 ELASTASE 2.
FT ACT_SITE 73 73 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 121 121 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 216 216 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 58 74 BY SIMILARITY.
FT DISULFID 155 222 BY SIMILARITY.
FT DISULFID 186 202 BY SIMILARITY.
FT DISULFID 212 243 BY SIMILARITY.
SQ SEQUENCE 269 AA; 28856 MW; 8343B97062CF267C CRC64;

Query Match 36.7%; Score 520; DB 1; Length 269;
Best Local Similarity 42.0%; Pred. No. 1.3e-37;
Matches 115; Conservative 45; Mismatches 98; Indels 16; Gaps 10;

QY 2 LLLSLTSLVLLGSSGCGVPAITPALSYNQRIVNCENAVPGSWPQVSLQDNTG---PH 58
Db 1 MIRALLSLVAG-ALSCGVPTYPPLS---RVVGEDARPNWPQVSLQYSSSCQWRH 56

QY 59 FCGSLISPNWVYTAACQVTPGRHVFVLGEXDRSSNAEPVQVLTARATHPHNNANTM 118
Db 57 TCGSLIEQNWLTAAHCISSTRTYRVVVGROSLFVSSESLTIAVSKSVIHEKWNQNL 116

QY 119 --NNDTLKLPARYTAQVSPVCLASTNEALPSGLTCVTGWGRISGVGNVTPARLQ 176
Db 117 AQGNDIALKLLASSVPLTDKIQGLCPAAGTILPNNVYCVYTWGRLQSG-ALPDILQ 175

QY 177 VVLPVTVNQCR--QYWGARITDAMICAGSG--ASSCGDSDGGLVCQKGN-TWVLIGIV 232
Db 176 KLLVVDYATCNSPWSWGVTYKTNMCAAGDGVTSNCNDSGGLNCAQANRQWVHGIV 235

QY 233 SMGTK-NCN-IQAPMYTRVSKFSTWVINOVMAYN 264
Db 236 SFGSSGLGCVYRKPSVTRVSNVNDWISSVIENN 269

RESULT 11
EL2_RAT

```

```

ID EL2_RAT STANDARD; PRT; 271 AA.
AC P00774;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Elastase 2 precursor (EC 3.4.21.71).
GN ELA2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82182967; PubMed=6918221;
RA McDonald R.J., Swift G.H., Quinto C., Swain W., Pictet R.L.,
RA Nikovits W., Rutter W.J.;
RT "Primary structure of two distinct rat pancreatic preproelastases
RT determined by sequence analysis of the complete cloned messenger
RT ribonucleic acid sequences.";
RL Biochemistry 21:1453-1463(1982).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85054882; PubMed=6094548;
RA Swift G.H., Craik C.S., Stary S.J., Quinto C., Lahaie R.G.,
RA Rutter W.J., Macdonald R.J.;
RT "Structure of the two related elastase genes expressed in the rat
RT pancreas.";
RL J. Biol. Chem. 259:14271-14278(1984).
CC -!- FUNCTION: ACTS UPON ELASTIN.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Leu|-Xaa, Met|-Xaa
CC and Phe|-Xaa. Hydrolyzes elastin.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: PANCREAS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; V01233; CAA24543.1; -.
DR EMBL; L00124; AAA98780.1; -.
DR EMBL; L00118; AAA98780.1; JOINED.
DR EMBL; L00119; AAA98780.1; JOINED.
DR EMBL; L00120; AAA98780.1; JOINED.
DR EMBL; L00121; AAA98780.1; JOINED.
DR EMBL; L00122; AAA98780.1; JOINED.
DR EMBL; L00123; AAA98780.1; JOINED.
DR PIR; A00961; ELRT2.
DR HSP; P00772; IELG.
DR MEROPS; S01.155; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS0240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolase; Serine protease; Zymogen; Signal.
FT SIGNAL 1 16
FT PROPEP 17 30 ACTIVATION PEPTIDE.
FT CHAIN 31 271 ELASTASE 2.
FT ACT_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 123 123 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 218 218 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 60 76 BY SIMILARITY.
FT DISULFID 157 224 BY SIMILARITY.
FT DISULFID 188 204 BY SIMILARITY.
FT DISULFID 214 245 BY SIMILARITY.

```

```
SQ SEQUENCE 271 AA; 28885 MW; 125C783B857B71E3 CRC64;
Query Match 36.58; Score 517; DB 1; Length 271;
Best Local Similarity 41.18; Pred. No. 2.4e-37;
Matches 113; Conservative 46; Mismatches 100; Indels 16; Gaps 11;

Qy 2 LLLSLLSLVLLGSSWGGVPAITPALSYN-ORVINGENAVPGSPWQVSLQ---DNTGF 57
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1 MIRTLLLSLVAG-ALSCGYP-IYEQHDYSRVVGGEASPNWQVSLQYLSSGKWH 57
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 58 HFCGSLISPNWVYTAHCQVTPGRHFVLLGEYDRSSNAEPVQVLSIARATHPNWNT 117
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 58 HFCGSLVANNWVLTAAHCISNSRTYRLLGRHSLSTSESGSLAVQVSKLVVHEKNAQK 117
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 118 MN--NDLTLLKLSPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQ 175
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 118 LSGNDIALVKLSPVALTSKIQTACLPAGTILPNNYPCVVTGMRGLQTVNG-ATPDVLQ 176
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 176 QVVLPLVTVNQC--ROYNGARITDAMICAGSG-ASSCGDGGPLVQKGN-TWVLIGI 231
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 177 QGRLLVVDYATCSSASWGGSSVKTMMVCAGDGVTTSSCGDGGPLNCOASNGQWQVHGI 236
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 232 VSWG-TKNCNI-QAPAMTRYVSKSTWVNOVMAYN 264
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 237 VSGSTLGCNYPKRPSPVETRVSNYIDWINSIAKN 271
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 12
EL2A_HUMAN STANDARD; PRT; 269 AA.
ID EL2A_HUMAN
AC P08217; Q14243;
DC 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE Elastase 2A precursor (EC 3.4.21.71).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=87217962; PubMed=3646943;
RA Kawashima I., Tani T., Shimoda K., Takiguchi Y.;
RT "Characterization of pancreatic elastase II cDNAs: two elastase II
RT mRNAs are expressed in human pancreas.";
RL DNA 6.163-172(1987).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=88107669; PubMed=3427074;
RA Fletcher T.S., Shen W.F., Largman C.;
RT "Primary structure of human pancreatic elastase 2 determined by
RT sequence analysis of the cloned mRNA.";
RL Biochemistry 26:7256-7261(1987).
RN [3]
RP SEQUENCE FROM N.A.
RA TISSUE=Pancreas;
RC MEDLINE=88198076; PubMed=2834346;
RA Shiraou Y., Yoshida H., Matsuki S., Takemura K., Ikeda N., Shimada Y.,
RA Ozawa T., Mikayama T., Iijima H., Ishida A., Sato Y., Tamai Y.,
RA Tanaka J., Ikenaga H.;
RT "Molecular cloning and expression in Escherichia coli of a cDNA
RT encoding human pancreatic elastase 2.";
RL J. Biochem. 102:1555-1563(1987).
RN [4]
RP SEQUENCE FROM N.A.
RA Thomas D.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA TISSUE=Pancreas;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ACTS UPON ELASTIN.
```

```
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Leu-|-Xaa, Met-|-Xaa
CC and Phe-|-Xaa. Hydrolyzes elastin.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: PANCREAS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M16631; AAA52374.1; -
CC EMBL; M16652; AAA52380.1; -
CC EMBL; D00236; BAA00165.1; -
CC EMBL; AL512883; CAC42421.1; -
CC EMBL; BC007031; AAH07031.1; -
CC PIR; A27432; A27432.
CC PIR; B26823; B26823.
CC HSP; P00772; IELG.
CC MEROPS; S01.155; -.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC SMART; SM00020; TRYP-SFC; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC KW Hydrolase; Serine protease; Zymogen; Signal.
CC FT SIGNAL 1 16
CC FT PROPEP 17 28 ACTIVATION PEPTIDE.
CC FT CHAIN 29 269 ELASTASE 2A.
CC FT DISULFID 58 74 BY SIMILARITY.
CC FT DISULFID 155 222 BY SIMILARITY.
CC FT DISULFID 186 202 BY SIMILARITY.
CC FT DISULFID 212 243 BY SIMILARITY.
CC FT ACT_SITE 73 73 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT ACT_SITE 121 121 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT ACT_SITE 216 216 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT CONFLICT 202 202 C -> V (IN REF. 3).
CC SQ SEQUENCE 269 AA; 28888 MW; A2E05143EFF4987C CRC64;

Query Match 36.38; Score 515; DB 1; Length 269;
Best Local Similarity 41.68; Pred. No. 3.5e-37;
Matches 114; Conservative 47; Mismatches 97; Indels 16; Gaps 10;

Qy 2 LLLSLLSLVLLGSSWGGVPAITPALSYNORVINGENAVPGSPWQVSLQDNTG---FH 58
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1 MIRTLLLSLVAG-ALSCGDPYPP---YVTRVVGGEARPNWQVSLQYSSNGKWH 56
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 59 FCGSLISPNWVYTAHCQVTPGRHFVLLGEYDRSSNAEPVQVLSIARATHPNWNT 118
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 57 TCGSLIANSWVLTAAHCISNSRTYRVLGRHSLVVAESGSLAVSVSKIVVHKDWSNOI 116
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 119 N--NDLTLLKLSPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQ 176
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 117 SKGNDIALVKLSPVSLTDKIQLACLPAGTILPNNYPCVVTGMRGLQTVNGAV-PDVLOQ 175
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 177 VVLPLVTVNQCQ--YWGARITDAMICAGSGA--SSCGDGGPLVQKGN-TWVLIGIV 232
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 176 GRLLVVDYATCSSASWGGSSVKTSMICAGDGVVSCNGDGGPLNCOASDGRWQVHGI 235
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 233 SWGK-NCN-IQAPAMTRYVSKSTWVNOVMAYN 264
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 236 SFGSLGCVNYHKPSVETRVSNYIDWINSVIAN 269
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 13
EL3B_HUMAN STANDARD; PRT; 270 AA.
ID EL3B_HUMAN
```

P08861; P11423;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Elastase IIIB precursor (EC 3.4.21.70) (Protease E).
 GN ELA3B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=88087253; PubMed=2826474;
 RA Tani T., Ohsumi J., Mita K., Takiguchi Y.;
 RT "Identification of a novel class of elastase isozyme, human
 RT pancreatic elastase III, by cDNA and genomic gene cloning.";
 RL J. Biol. Chem. 263:1231-1239(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 4-270 FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=88000845; PubMed=3477287;
 RA Shen W., Fletcher T.S., Largman C.;
 RT "Primary structure of human pancreatic protease E determined by
 RT sequence analysis of the cloned mRNA.";
 RL Biochemistry 26:3447-3452(1987).
 RN [4]
 RP SEQUENCE OF 31-50.
 RC TISSUE=Pancreas;
 RX MEDLINE=89325560; PubMed=2753124;
 RA Moulard M., Kerfelec B., Mallet B., Chapus C.;
 RT "Identification of a procarboxypeptidase A-truncated protease E
 RT binary complex in human pancreatic juice.";
 RL FEBS Lett. 250:166-170(1989).
 RN [5]
 RP SEQUENCE OF 94-164, AND CARBOHYDRATE-LINKAGE SITES ASN-114.
 RC TISSUE=Pancreas;
 RX MEDLINE=89289996; PubMed=2737288;
 RA Wendorf P., Geyer R., Szlegoleit A., Linder D.;
 RT "Localization and characterization of the glycosylation site of human
 RT pancreatic elastase 1.";
 RL FEBS Lett. 249:275-278(1989).
 CC -!- FUNCTION: EFFICIENT PROTEASE WITH ALANINE SPECIFICITY BUT ONLY
 CC LITTLE ELASTOLYTIC ACTIVITY.
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Ala-I-Xaa. Does not
 CC hydrolyse elastin.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
 CC -!- CAUTION: Was originally (Ref.5) thought to be elastase 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M16630; AAA36482.1; -
 CC DR EMBL; BC005216; AAH05216.1; -
 CC DR EMBL; M18692; AAA58434.1; -
 CC DR PIR; B29934; B29934.
 CC DR PIR; A27206; A27206.
 CC DR PIR; S04999; S04999.
 CC DR PIR; S04490; S04490.
 CC DR HSP; P05805; 1FON.
 CC DR MEROPS; S01.205; -
 CC DR GlycosuiteBB; P08861; -
 CC DR SWISS-2DPAGE; P08861; HUMAN.

DR Genew; HGNC:15945; ELA3B.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PRO0722; CHYMOTRYPsin.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Zymogen; Signal; Glycoprotein.
 FT SIGNAL 1 15
 FT PROPEP 16 28 OR 16 (POTENTIAL).
 FT CHAIN 29 270 ACTIVATION PEPTIDE (POTENTIAL).
 FT ACT_SITE 73 73 ELASTASE IIIB.
 FT ACT_SITE 123 123 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 217 217 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 58 74 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 117 120 PROBABLE.
 FT DISULFID 157 223 BY SIMILARITY.
 FT DISULFID 188 204 BY SIMILARITY.
 FT DISULFID 213 244 BY SIMILARITY.
 FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .).
 FT CONFLICT 4 4 /FTIQ-CAR_000212.
 FT CONFLICT 64 64 R -> G (IN REF. 3).
 FT CONFLICT 79 79 A -> G (IN REF. 3).
 FT CONFLICT 129 131 W -> R (IN REF. 1).
 FT CONFLICT 164 164 MISSING (IN REF. 5).
 FT CONFLICT 184 184 R -> P (IN REF. 3).
 SQ SEQUENCE 270 AA; 29293 MW; B14BE0AAD369SAFE CRC64;
 Query Match 36.2%; Score 513.5; DB 1; Length 270;
 Best Local Similarity 41.4%; Pred. No. 4.7e-37;
 Matches 113; Conservative 46; Mismatches 97; Indels 17; Gaps 10;
 QY 2 LLSLTSLVLLGSSWCGVPAITPALSPALSNQRIIVNGENAVPGSWPQVSLQ-DNTG--FH 58
 DB 1 MMLRLSSLLVAVASGYGPPSSRP-----SSRVNGEDAVPYSPWPQVSLQYKSGSFYH 56
 QY 59 FCGGSLISPNWVVTAAHCQVTPGRHFVVLGEYDRSSNAEPVQVLSI--ARAITHPNWNA 116
 DB 57 TCGSLIADPWVVTAGHCISSSWTYQVVLGEYDRAVKEGPEQVPIPSNGDLFVHPLNRS 116
 QY 117 TM--NNDLTLKLASPARYTAQVSPVCLASTNEALPSGLTCVTWGTGRISGVGNVT PARL 174
 DB 117 CVACGNIDIALKLSRAQLGDAVOLASLPAGDILNETPCYITGWRGLYTNGL-PDKL 175
 QY 175 QQVVLPLTVNOCRO--YWGARTIDAMICAGSGCSCQDGGPLVC-QKGNTWVLI GI 231
 DB 176 QEALLPVDYEHCSRWNWNGSSVYKTMVCAGGDIRSGCNGDGGPLNCPTEGGVQVHGV 235
 QY 232 VSW-GFKNCNI-QAPAMYTRVSKFSTWVQVMA 262
 DB 236 TSEVSAFGCNTRKPKVFTVRSFAFDWIBETIA 268
 RESULT 14
 CAC3 BOVIN
 ID CAC3 BOVIN STANDARD; PRT; 253 AA.
 AC P05805;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Proproteinase E precursor (procarboxypeptidase A complex component
 DE III) (Procarboxypeptidase A-S6 subunit III) (PROCPA-S6 III).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE OF 1-25.
 RX MEDLINE=91099520; PubMed=22693366;
 RA Pascual R., Vendrell J., Aviles F.X., Bonicel J., Wicker C.,

Search completed: December 20, 2002, 15:17:13
Job time : 10.0435 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:14:28 : Search time 16.6957 Seconds
(without alignments)
1520.126 Million cell updates/sec

Title: US-09-856-319b-4
Perfect score: 1418
Sequence: 1 MLLSLTLTLVLLGSGCG.....AMVTRYSKFTWVQVAYN 264

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR-73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1269	89.5	264	2 I38136	chymotrypsin-like
2	832.5	58.7	263	2 A21195	chymotrypsin (EC 3
3	822.5	58.0	263	2 A31299	chymotrypsin (EC 3
4	798.5	56.3	263	2 S47537	chymotrypsin (EC 3
5	786.5	55.5	245	1 KYBOB	chymotrypsin (EC 3
6	781.5	55.1	263	1 KYRTB	chymotrypsin (EC 3
7	774.5	54.6	245	1 KYBOA	chymotrypsin (EC 3
8	737	52.0	244	2 S72219	chymotrypsin B - A
9	537	37.9	269	2 A26823	pancreatic elastase
10	517	36.5	271	1 ELRT2	pancreatic elastase
11	515	36.3	269	2 B26823	pancreatic elastase
12	513.5	36.2	270	2 B29934	pancreatic elastase
13	499	35.2	1524	2 T30337	polyprotein - Afri
14	495.5	34.9	268	2 S68826	pancreatic elastase
15	495.5	34.9	268	2 S68825	pancreatic elastase
16	492	34.7	271	2 A25528	pancreatic elastase
17	485.5	34.2	240	1 CPBOA3	procarboxypeptidase
18	485.5	34.2	270	2 A29934	pancreatic elastase
19	485.5	34.2	812	1 PLMS	plasmin (EC 3.4.21
20	484	34.1	810	1 PLHU	plasmin (EC 3.4.21
21	479	33.8	269	2 C26823	pancreatic elastase
22	477	33.6	126	2 A23473	chymotrypsin-like
23	477	33.6	460	2 B61545	plasmin (EC 3.4.21
24	475	33.5	810	2 B30848	plasmin (EC 3.4.21
25	474	33.4	790	1 PLPG	plasmin (EC 3.4.21
26	465	32.8	270	2 S56160	mast cell tryptase
27	463.5	32.7	273	2 A47246	tryptase (EC 3.4.2
28	462.5	32.6	343	1 A57014	proctasin (EC 3.4.
29	459	32.4	455	2 A61545	plasmin (EC 3.4.21

30 454.5 32.1 246 1 TRRT1
31 453.5 32.0 268 2 JQ1473
32 447 31.5 231 1 TRPGPR
33 445.5 31.4 274 2 JC4171
34 444 31.3 266 1 ELRT1
35 443.5 31.3 246 1 TRRT2
36 443.5 31.3 275 2 A32410
37 443 31.2 810 2 I46260
38 442.5 31.2 246 1 TRDGC
39 441.5 31.1 812 1 PLBO
40 441 31.1 786 1 A47547
41 440.5 31.1 247 1 TRDG
42 440.5 31.1 276 2 A38654
43 440.5 31.1 4548 1 S00657
44 439.5 31.0 1035 1 A43090
45 439 31.0 1420 2 A32869

ALIGNMENTS

RESULT 1
I38136
chymotrypsin-like proteinase (EC 3.4.21.-) CTRL-1 - human
C:Species: Homo sapiens (man)
C:Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 22-Jun-1999
C:Accession: I38136
R:Larsen, F.; Solheim, J.; Kristensen, T.; Kolsto, A.B.; Prydz, H.
Hum. Mol. Genet. 2, 1589-1595, 1993
A:Title: A tight cluster of five unrelated human genes on chromosome 16q22.1.
A:Reference number: I38135; MUID:94093544; PMID:8268911
A:Accession: I38136
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-264 <RES>
A:Cross-references: EMBL:X71874; NID:g406226; PIDN:CAA50710.1; PID:g406228
C:Genetics:
A:Gene: GDB:CTRL
A:Cross-references: GDB:204061
A:Map position: 16q22.1-16q22.1
A:Introns: 18/1; 52/3; 79/2; 106/3; 167/1; 211/3
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:34-257/Domain: trypsin homology <TRY>
F:75,121,214/Active Site: His, Asp, Ser #status predicted

Query Match 89.5%; Score 1269; DB 2; Length 264;
Best Local Similarity 86.4%; Pred. No. 9.9e-98;
Matches 228; Conservative 25; Mismatches 11; Indels 0; Gaps 0;

Qy 1 MLLSLTLTLVLLGSGCGVPAITPALSYNQIRVNGENAVPGSWPQVSLQDNTGEHFC 60
Db 1 MLLSLTLTLVLLGSGCGGIPAIKPAUSFSQIRVNGENAVLGSPWQVSLQDSSGFHFC 60
Qy 61 GGSLSIPNVVYTAACQVTPGRHFHVLGEYDRSSNAEPQVLSARATHPNWNTMNN 120
Db 61 GGSLSISQSVVYTAACNVSPGRHFHVLGEYDRSSNAEPQLQVLSVSRATHTPSWNTMNN 120
Qy 121 DLTLLKASPARYAQVSPVCLASTNEALPSGLTCVTTGWGRISGVNVTPARLQOVVLP 180
Db 121 DVTLLKASPAQYTRISPVCLASNEALTEGLTCVTTGWGRISGVNVTPARLQOVVLP 180
Qy 181 LVTYNQCRQYWGARTDMICAGSGASSCGDGGPLVCOKGNTWVLLIGVSGTKNCN 240
Db 181 LVTYNQCRQYWGSSITDSMICAGGAGASSCGDGGPLVCOKGNTWVLLIGVSGTKNCN 240
Qy 241 IQAPAMVTRVSKFTWVQVAYN 264
Db 241 VRAPVTRVSKFTWVQVAYN 264

RESULT 2
A21195

chymotrypsin (EC 3.4.21.1) 2 precursor - dog
 C:Species: Canis lupus familiaris (dog)
 C>Date: 27-Jul-1990 #sequence_revision 22-Jul-1990 #text_change 22-Jun-1999
 C:Accession: A21195
 R:Pinsky, S.D.; LaForge, K.S.; Luc, V.; Scheele, G.
 Proc. Natl. Acad. Sci. U.S.A. 80, 7486-7490, 1983
 A:Title: Identification of cDNA clones encoding secretory isoenzyme forms: sequence details
 A:Reference number: A21195; MUID:84170253; PMID:6584866
 A:Accession: A21195
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-263 <PIN>
 A:Cross-references: GB:K01173; NID:g163945; PIDN:AAA30841.1; PID:g163946
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; protein digestion; serine proteinase
 F:34-256/Domain: trypsin homology <TRY>
 F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 58.7%; Score 832.5; DB 2; Length 263;
 Best Local Similarity 57.2%; Pred. No. 1.3e-61;
 Matches 151; Conservative 38; Mismatches 74; Indels 1; Gaps 1;

QY 1 MLLSLTSLVLGSGWCGVPAITPALSINORIVNGENAVPGSWPQVSLQDNTGFGHFC 60
 DB 1 MAFLLWLLSCFALLGTAFGCGVPAIQVLSGLSRIVNGEDAVPGSWPQVSLQDSTGFGHFC 60
 QY 61 GGSLSIPNNVWVTAACHQVTPGRHFVVLGVDYDRSSNAEPVQVLSIARAIHPNNANTMNN 120
 DB 61 GGSLSIEDWVTAACHQVTPGRHFVVLGVDYDRSSNAEPVQVLSIARAIHPNNANTMNN 119
 QY 121 DITLLKSLAPARYTAQVSPVCLASTNEALPSGLTCTVTGWRISGVGNVTPARLQOVVLP 180
 DB 121 DITLLKSLAPARYTAQVSPVCLASTNEALPSGLTCTVTGWRISGVGNVTPARLQOVVLP 179
 QY 181 LVTNOCROYNGARITDAMICAGGSGASSCGSGPLVCQKGNWTWVLGIVSWGKNCN 240
 DB 181 LVTNOCROYNGARITDAMICAGGSGASSCGSGPLVCQKGNWTWVLGIVSWGKNCN 239
 QY 241 IOAPAMYTRVSKFSTWVNOVMAYN 264
 DB 241 IOAPAMYTRVSKFSTWVNOVMAYN 263
 QY 263 TSPGVIARVTKLIPWVQILQAN 263
 DB 263 TSPGVIARVTKLIPWVQILQAN 263

RESULT 3
 chymotrypsin (EC 3.4.21.1) precursor - human
 C:Species: Homo sapiens (man)
 C>Date: 08-Jun-1989 #sequence_revision 08-Jun-1989 #text_change 22-Jun-1999
 C:Accession: A31299
 R:Tomita, N.; Izumoto, Y.; Horii, A.; Doi, S.; Yokouchi, H.; Ogawa, M.; Mori, T.; Matsubara, T.
 Biochem. Biophys. Res. Commun. 158, 569-575, 1989
 A:Title: Molecular cloning and nucleotide sequence of human pancreatic precchymotrypsinogen
 A:Reference number: A31299; MUID:89134264; PMID:2917002
 A:Accession: A31299
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-263 <TOM>
 A:Cross-references: GB:M24400; NID:g181189; PIDN:AAA52128.1; PID:g181190
 C:Genetics:
 A:Gene: GDB:CTRB1; CTRB
 A:Cross-references: GDB:119820; OMIM:118890
 A:Map position: 16q23.1-16q23.1
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; protein digestion; serine proteinase
 F:34-256/Domain: trypsin homology <TRY>
 F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 58.0%; Score 822.5; DB 2; Length 263;
 Best Local Similarity 57.2%; Pred. No. 9.1e-61;
 Matches 151; Conservative 38; Mismatches 74; Indels 1; Gaps 1;

QY 1 MLLSLTSLVLGSGWCGVPAITPALSINORIVNGENAVPGSWPQVSLQDNTGFGHFC 60
 DB 1 MAFLLWLLSCFALLGTAFGCGVPAIQVLSGLSRIVNGEDAVPGSWPQVSLQDSTGFGHFC 60

Db 1 MAFLLWLLSCWALLGTTFGCGVPAIHVLSGLSRIVNGEDAVPGSWPQVSLQDNTGFGHFC 60
 QY 61 GGSLSIPNNVWVTAACHQVTPGRHFVVLGVDYDRSSNAEPVQVLSIARAIHPNNANTMNN 120
 DB 61 GGSLSIEDWVTAACHQVTPGRHFVVLGVDYDRSSNAEPVQVLSIARAIHPNNANTMNN 119
 QY 121 DITLLKSLAPARYTAQVSPVCLASTNEALPSGLTCTVTGWRISGVGNVTPARLQOVVLP 180
 DB 121 DITLLKSLAPARYTAQVSPVCLASTNEALPSGLTCTVTGWRISGVGNVTPARLQOVVLP 179
 QY 181 LVTNOCROYNGARITDAMICAGGSGASSCGSGPLVCQKGNWTWVLGIVSWGKNCN 240
 DB 181 LVTNOCROYNGARITDAMICAGGSGASSCGSGPLVCQKGNWTWVLGIVSWGKNCN 239
 QY 241 IOAPAMYTRVSKFSTWVNOVMAYN 264
 DB 241 IOAPAMYTRVSKFSTWVNOVMAYN 263
 QY 263 TSPGVIARVTKLIPWVQILQAN 263
 DB 263 TSPGVIARVTKLIPWVQILQAN 263

RESULT 4
 S47537
 chymotrypsin (EC 3.4.21.1) precursor - Atlantic cod
 C:Species: Gadus morhua (Atlantic cod)
 C>Date: 26-Dec-1994 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999
 C:Accession: S47537; S43163
 R:Gudmundsdottir, A.; Oskarsson, S.; Eakin, A.E.; Craik, C.S.; Bjarnason, J.B.
 Biochim. Biophys. Acta 1219, 211-214, 1994
 A:Title: Atlantic cod cDNA encoding a psychrophilic chymotrypsinogen.
 A:Reference number: S47537; MUID:94368860; PMID:8086467
 A:Accession: S47537
 A:Molecule type: mRNA
 A:Residues: 1-263 <GUD>
 A:Cross-references: EMBL:X87490; NID:g468750; PIDN:CAA55242.1; PID:g468751
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; protein digestion; serine proteinase
 F:1-18/Domain: signal sequence #status predicted <Sig>
 F:19-263/Product: chymotrypsin #status predicted <MAT>
 F:34-256/Domain: trypsin homology <TRY>
 F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 56.3%; Score 798.5; DB 2; Length 263;
 Best Local Similarity 56.9%; Pred. No. 8.8e-59;
 Matches 144; Conservative 38; Mismatches 70; Indels 1; Gaps 1;

QY 12 LILSGWCGVPAITPALSINORIVNGENAVPGSWPQVSLQDNTGFGHFCGSLIPNNVWV 71
 DB 12 LFRRTYCGGRPAISPVITGYSRIVNGEAVPHSWNQVSLQDQGTGFGHFCGSLIPNNVWV 71
 QY 72 TAAHQVTPGRHFVVLGVDYDRSSNAEPVQVLSIARAIHPNNANTMNDLLKSLAP 131
 DB 72 TAAHQVTPGRHFVVLGVDYDRSSNAEPVQVLSIARAIHPNNANTMNDLLKSLAP 130
 QY 132 RYTAQVSPVCLASTNEALPSGLTCTVTGWRISGVGNVTPARLQOVVLPVNOCRQW 191
 DB 132 RYTAQVSPVCLASTNEALPSGLTCTVTGWRISGVGNVTPARLQOVVLPVNOCRQW 190
 QY 192 GARITDAMICAGGSGASSCGSGPLVCQKGNWTWVLGIVSWGKNCNIPAMYTRYS 251
 DB 192 GARITDAMICAGGSGASSCGSGPLVCQKGNWTWVLGIVSWGKNCNIPAMYTRYS 250
 QY 252 KFTWINQVMAYN 264
 DB 252 KFTWINQVMAYN 263
 QY 263 TSPGVIARVTKLIPWVQILQAN 263
 DB 263 TSPGVIARVTKLIPWVQILQAN 263

RESULT 5
 K7B0B
 chymotrypsin (EC 3.4.21.1) B precursor - bovine
 N:Alternate names: chymotrypsinogen B
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 08-Oct-1981 #sequence_revision 08-Oct-1981 #text_change 18-Jul-1997
 C:Accession: A00953
 R:Smillie, L.B.; Furka, A.; Nagabhushan, N.; Stevenson, K.J.; Parkes, C.O.

Nature 218, 343-346, 1968
A:Title: Structure of chymotrypsinogen B compared with chymotrypsinogen A and trypsinogen
A:Reference number: A00953; MUID:68238908; PMID:5649671
A:Accession: A00953
A:Molecule type: protein
A:Residues: 1-245 <SMI>
C:Comment: Chymotrypsinogen B is synthesized, along with chymotrypsinogen A, in the acinar cells of the pancreas. The first activation cleavage, leading to pI-chymotrypsin B, occurs in the secretory granules of the acinar cells.
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolyase; pancreas; protein digestion; serine proteinase; zymogen
F:1-15/Domain: propeptide #status experimental <PRO>
F:16-245/Product: chymotrypsin B #status experimental <MAT>
F:16-238/Domain: trypsin homology <TRY>
F:1-122,42-58,136-201,168-182,191-220/Disulfide bonds: #status experimental
F:57,102,195/Active site: His, Asp, Ser #status experimental

Query Match 55.5%; Score 786.5; DB 1; Length 245;
Best Local Similarity 57.7%; Pred. No. 8e-58;
Matches 142; Conservative 32; Mismatches 71; Indels 1; Gaps 1;

Qy 19 CGVPAITPALSIVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVVTAAHCOV 78
Db 1 CGVPAIQPVLSGLARIVNGEDAVPGSWPQVSLQDSTGFHFCGSLISEDWVVTAAHCOV 60

Qy 79 TPGRFHVVLGYDRSSNAEPVQVLSIARAIHPNNANTMNDLTLKLASPARYTAQVS 138
Db 61 TTS-DVVVAGEFDGLETDQVLKIGKVFKNPKFSILTVRNDITLLKLATPAQFSETVS 119

Qy 139 PVCLASTNEALPSGLTCVTGHRISGVGNVTPARLQVPLVTVNQCQVWGAIRYDA 198
Db 120 AVCLPSADEDFPAGMLCATGTGKTKYNALKTPTDQLQOATLPIVNTDCRKWGSRVTDV 179

Qy 199 MICAGSGASSCGDGGPLVCQKNTWVLIGVSWGKNCNIQAPAMTVTRVSKFSTWIN 258
Db 180 MICAGASGVSSCMGSGGGLVCCOKNGAWTLGIVSWGSGTCTSTPVAIVARVTLMPWQV 239

Qy 259 QVMAYN 264
Db 240 ETLAN 245

RESULT 6
KYTB
chymotrypsin (EC 3.4.21.1) B precursor - rat
N:Alternate names: chymotrypsinogen B
C:Species: Rattus norvegicus (Norway rat)
C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 18-Jun-1999
C:Accession: A22658
R:Beil, G.I.; Quinto, C.; Quiroga, M.; Valenzuela, P.; Craik, C.S.; Rutter, W.J.
J. Biol. Chem. 259, 14265-14270, 1984
A:Title: Isolation and sequence of a rat chymotrypsin B gene.
A:Reference number: A22658; MUID:85054881; PMID:6209274
A:Accession: A22658
A:Molecule type: DNA
A:Residues: 1-263 <BEL>
A:Cross-references: GB:K02298; NID:g203653; PIDN:AAA98732.1; PID:g203654
C:Introns: 18/1; 52/3; 79/2; 105/3; 166/1; 210/3
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolyase; pancreas; protein digestion; serine proteinase
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-33/Domain: propeptide #status predicted <PRO>
F:34-263/Product: chymotrypsin B #status predicted <MAT>
F:34-256/Domain: trypsin homology <TRY>
F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 55.1%; Score 781.5; DB 1; Length 263;
Best Local Similarity 53.8%; Pred. No. 2.2e-57;
Matches 142; Conservative 40; Mismatches 81; Indels 1; Gaps 1;

Qy 1 MULLSLTSLVLGSGWCGVPAITPALSIVNGENAVPGSWPQVSLQDNTGFHFC 60
Db 1 MAPLWLVSCFALVGATFGCGVPTIQVLTLGLSRVNGEDAIPEGSPWQVSLQDKTGFHFC 60

Qy 61 GGSLLSPNNVVTAAHCOVTPGRHFRVVLGYDRSSNAEPVQVLSIARAIHPNNANTMNN 120
Db 61 GGSLLSEDWVVTAAHCGVKTS-DVVVAGEFDQGSDEENIQVLKIAQVFNKFNFTVRN 119

Qy 121 DUTLLKLASPARYTAQVSPVCLASTNEALPSGLTCVTGHRISGVGNVTPARLQVPLP 180
Db 120 DITLLKLATPAQFSETVSACLPNVDDDFPCTVCATTGKTKYNALKTPEKLOQAALP 179

Qy 181 LTVNQCQVWGAIRYTDAMICAGSGASSCGDGGPLVCQKNTWVLIGVSWGKNCN 240
Db 180 IVSEADCKKSGSKITDVTMTACASGVSSCMGSGGGLVCCOKDGVWTLGIVSWGSGVCS 239

Qy 241 IQAPAMTVTRVSKFSTWINQVMAYN 264
Db 240 TSTPAVYSRVLTALMPWQVQILEAN 263

RESULT 7
KYBOA
chymotrypsin (EC 3.4.21.1) A precursor - bovine
N:Alternate names: chymotrypsinogen A
C:Species: Bos primigenius taurus (cattle)
C:Date: 07-May-1981 #sequence_revision 07-May-1981 #text_change 07-May-1999
C:Accession: A30235; A93158; S29650; A00952
R:Brown, J.R.; Hartley, B.S.
Biochem. J. 101, 214-228, 1966
A:Title: Location of disulphide bridges by diagonal paper electrophoresis. The disulphide bridges are located at positions 101-102, 103-104, 105-106, 107-108, 109-110, 111-112, 113-114, 115-116, 117-118, 119-120, 121-122, 123-124, 125-126, 127-128, 129-130, 131-132, 133-134, 135-136, 137-138, 139-140, 141-142, 143-144, 145-146, 147-148, 149-150, 151-152, 153-154, 155-156, 157-158, 159-160, 161-162, 163-164, 165-166, 167-168, 169-170, 171-172, 173-174, 175-176, 177-178, 179-180, 181-182, 183-184, 185-186, 187-188, 189-190, 191-192, 193-194, 195-196, 197-198, 199-200, 201-202, 203-204, 205-206, 207-208, 209-210, 211-212, 213-214, 215-216, 217-218, 219-220, 221-222, 223-224, 225-226, 227-228, 229-230, 231-232, 233-234, 235-236, 237-238, 239-240, 241-242, 243-244, 245-246, 247-248, 249-250, 251-252, 253-254, 255-256, 257-258, 259-260, 261-262, 263-264, 265-266, 267-268, 269-270, 271-272, 273-274, 275-276, 277-278, 279-280, 281-282, 283-284, 285-286, 287-288, 289-290, 291-292, 293-294, 295-296, 297-298, 299-300, 301-302, 303-304, 305-306, 307-308, 309-310, 311-312, 313-314, 315-316, 317-318, 319-320, 321-322, 323-324, 325-326, 327-328, 329-330, 331-332, 333-334, 335-336, 337-338, 339-340, 341-342, 343-344, 345-346, 347-348, 349-350, 351-352, 353-354, 355-356, 357-358, 359-360, 361-362, 363-364, 365-366, 367-368, 369-370, 371-372, 373-374, 375-376, 377-378, 379-380, 381-382, 383-384, 385-386, 387-388, 389-390, 391-392, 393-394, 395-396, 397-398, 399-400, 401-402, 403-404, 405-406, 407-408, 409-410, 411-412, 413-414, 415-416, 417-418, 419-420, 421-422, 423-424, 425-426, 427-428, 429-430, 431-432, 433-434, 435-436, 437-438, 439-440, 441-442, 443-444, 445-446, 447-448, 449-450, 451-452, 453-454, 455-456, 457-458, 459-460, 461-462, 463-464, 465-466, 467-468, 469-470, 471-472, 473-474, 475-476, 477-478, 479-480, 481-482, 483-484, 485-486, 487-488, 489-490, 491-492, 493-494, 495-496, 497-498, 499-500, 501-502, 503-504, 505-506, 507-508, 509-510, 511-512, 513-514, 515-516, 517-518, 519-520, 521-522, 523-524, 525-526, 527-528, 529-530, 531-532, 533-534, 535-536, 537-538, 539-540, 541-542, 543-544, 545-546, 547-548, 549-550, 551-552, 553-554, 555-556, 557-558, 559-560, 561-562, 563-564, 565-566, 567-568, 569-570, 571-572, 573-574, 575-576, 577-578, 579-580, 581-582, 583-584, 585-586, 587-588, 589-590, 591-592, 593-594, 595-596, 597-598, 599-600, 601-602, 603-604, 605-606, 607-608, 609-610, 611-612, 613-614, 615-616, 617-618, 619-620, 621-622, 623-624, 625-626, 627-628, 629-630, 631-632, 633-634, 635-636, 637-638, 639-640, 641-642, 643-644, 645-646, 647-648, 649-650, 651-652, 653-654, 655-656, 657-658, 659-660, 661-662, 663-664, 665-666, 667-668, 669-670, 671-672, 673-674, 675-676, 677-678, 679-680, 681-682, 683-684, 685-686, 687-688, 689-690, 691-692, 693-694, 695-696, 697-698, 699-700, 701-702, 703-704, 705-706, 707-708, 709-710, 711-712, 713-714, 715-716, 717-718, 719-720, 721-722, 723-724, 725-726, 727-728, 729-730, 731-732, 733-734, 735-736, 737-738, 739-740, 741-742, 743-744, 745-746, 747-748, 749-750, 751-752, 753-754, 755-756, 757-758, 759-760, 761-762, 763-764, 765-766, 767-768, 769-770, 771-772, 773-774, 775-776, 777-778, 779-780, 781-782, 783-784, 785-786, 787-788, 789-790, 791-792, 793-794, 795-796, 797-798, 799-800, 801-802, 803-804, 805-806, 807-808, 809-810, 811-812, 813-814, 815-816, 817-818, 819-820, 821-822, 823-824, 825-826, 827-828, 829-830, 831-832, 833-834, 835-836, 837-838, 839-840, 841-842, 843-844, 845-846, 847-848, 849-850, 851-852, 853-854, 855-856, 857-858, 859-860, 861-862, 863-864, 865-866, 867-868, 869-870, 871-872, 873-874, 875-876, 877-878, 879-880, 881-882, 883-884, 885-886, 887-888, 889-890, 891-892, 893-894, 895-896, 897-898, 899-900, 901-902, 903-904, 905-906, 907-908, 909-910, 911-912, 913-914, 915-916, 917-918, 919-920, 921-922, 923-924, 925-926, 927-928, 929-930, 931-932, 933-934, 935-936, 937-938, 939-940, 941-942, 943-944, 945-946, 947-948, 949-950, 951-952, 953-954, 955-956, 957-958, 959-960, 961-962, 963-964, 965-966, 967-968, 969-970, 971-972, 973-974, 975-976, 977-978, 979-980, 981-982, 983-984, 985-986, 987-988, 989-990, 991-992, 993-994, 995-996, 997-998, 999-1000, 1001-1002, 1003-1004, 1005-1006, 1007-1008, 1009-1010, 1011-1012, 1013-1014, 1015-1016, 1017-1018, 1019-1020, 1021-1022, 1023-1024, 1025-1026, 1027-1028, 1029-1030, 1031-1032, 1033-1034, 1035-1036, 1037-1038, 1039-1040, 1041-1042, 1043-1044, 1045-1046, 1047-1048, 1049-1050, 1051-1052, 1053-1054, 1055-1056, 1057-1058, 1059-1060, 1061-1062, 1063-1064, 1065-1066, 1067-1068, 1069-1070, 1071-1072, 1073-1074, 1075-1076, 1077-1078, 1079-1080, 1081-1082, 1083-1084, 1085-1086, 1087-1088, 1089-1090, 1091-1092, 1093-1094, 1095-1096, 1097-1098, 1099-1100, 1101-1102, 1103-1104, 1105-1106, 1107-1108, 1109-1110, 1111-1112, 1113-1114, 1115-1116, 1117-1118, 1119-1120, 1121-1122, 1123-1124, 1125-1126, 1127-1128, 1129-1130, 1131-1132, 1133-1134, 1135-1136, 1137-1138, 1139-1140, 1141-1142, 1143-1144, 1145-1146, 1147-1148, 1149-1150, 1151-1152, 1153-1154, 1155-1156, 1157-1158, 1159-1160, 1161-1162, 1163-1164, 1165-1166, 1167-1168, 1169-1170, 1171-1172, 1173-1174, 1175-1176, 1177-1178, 1179-1180, 1181-1182, 1183-1184, 1185-1186, 1187-1188, 1189-1190, 1191-1192, 1193-1194, 1195-1196, 1197-1198, 1199-1200, 1201-1202, 1203-1204, 1205-1206, 1207-1208, 1209-1210, 1211-1212, 1213-1214, 1215-1216, 1217-1218, 1219-1220, 1221-1222, 1223-1224, 1225-1226, 1227-1228, 1229-1230, 1231-1232, 1233-1234, 1235-1236, 1237-1238, 1239-1240, 1241-1242, 1243-1244, 1245-1246, 1247-1248, 1249-1250, 1251-1252, 1253-1254, 1255-1256, 1257-1258, 1259-1260, 1261-1262, 1263-1264, 1265-1266, 1267-1268, 1269-1270, 1271-1272, 1273-1274, 1275-1276, 1277-1278, 1279-1280, 1281-1282, 1283-1284, 1285-1286, 1287-1288, 1289-1290, 1291-1292, 1293-1294, 1295-1296, 1297-1298, 1299-1300, 1301-1302, 1303-1304, 1305-1306, 1307-1308, 1309-1310, 1311-1312, 1313-1314, 1315-1316, 1317-1318, 1319-1320, 1321-1322, 1323-1324, 1325-1326, 1327-1328, 1329-1330, 1331-1332, 1333-1334, 1335-1336, 1337-1338, 1339-1340, 1341-1342, 1343-1344, 1345-1346, 1347-1348, 1349-1350, 1351-1352, 1353-1354, 1355-1356, 1357-1358, 1359-1360, 1361-1362, 1363-1364, 1365-1366, 1367-1368, 1369-1370, 1371-1372, 1373-1374, 1375-1376, 1377-1378, 1379-1380, 1381-1382, 1383-1384, 1385-1386, 1387-1388, 1389-1390, 1391-1392, 1393-1394, 1395-1396, 1397-1398, 1399-1400, 1401-1402, 1403-1404, 1405-1406, 1407-1408, 1409-1410, 1411-1412, 1413-1414, 1415-1416, 1417-1418, 1419-1420, 1421-1422, 1423-1424, 1425-1426, 1427-1428, 1429-1430, 1431-1432, 1433-1434, 1435-1436, 1437-1438, 1439-1440, 1441-1442, 1443-1444, 1445-1446, 1447-1448, 1449-1450, 1451-1452, 1453-1454, 1455-1456, 1457-1458, 1459-1460, 1461-1462, 1463-1464, 1465-1466, 1467-1468, 1469-1470, 1471-1472, 1473-1474, 1475-1476, 1477-1478, 1479-1480, 1481-1482, 1483-1484, 1485-1486, 1487-1488, 1489-1490, 1491-1492, 1493-1494, 1495-1496, 1497-1498, 1499-1500, 1501-1502, 1503-1504, 1505-1506, 1507-1508, 1509-1510, 1511-1512, 1513-1514, 1515-1516, 1517-1518, 1519-1520, 1521-1522, 1523-1524, 1525-1526, 1527-1528, 1529-1530, 1531-1532, 1533-1534, 1535-1536, 1537-1538, 1539-1540, 1541-1542, 1543-1544, 1545-1546, 1547-1548, 1549-1550, 1551-1552, 1553-1554, 1555-1556, 1557-1558, 1559-1560, 1561-1562, 1563-1564, 1565-1566, 1567-1568, 1569-1570, 1571-1572, 1573-1574, 1575-1576, 1577-1578, 1579-1580, 1581-1582, 1583-1584, 1585-1586, 1587-1588, 1589-1590, 1591-1592, 1593-1594, 1595-1596, 1597-1598, 1599-1600, 1601-1602, 1603-1604, 1605-1606, 1607-1608, 1609-1610, 1611-1612, 1613-1614, 1615-1616, 1617-1618, 1619-1620, 1621-1622, 1623-1624, 1625-1626, 1627-1628, 1629-1630, 1631-1632, 1633-1634, 1635-1636, 1637-1638, 1639-1640, 1641-1642, 1643-1644, 1645-1646, 1647-1648, 1649-1650, 1651-1652, 1653-1654, 1655-1656, 1657-1658, 1659-1660, 1661-1662, 1663-1664, 1665-1666, 1667-1668, 1669-1670, 1671-1672, 1673-1674, 1675-1676, 1677-1678, 1679-1680, 1681-1682, 1683-1684, 1685-1686, 1687-1688, 1689-1690, 1691-1692, 1693-1694, 1695-1696, 1697-1698, 1699-1700, 1701-1702, 1703-1704, 1705-1706, 1707-1708, 1709-1710, 1711-1712, 1713-1714, 1715-1716, 1717-1718, 1719-1720, 1721-1722, 1723-1724, 1725-1726, 1727-1728, 1729-1730, 1731-1732, 1733-1734, 1735-1736, 1737-1738, 1739-1740, 1741-1742, 1743-1744, 1745-1746, 1747-1748, 1749-1750, 1751-1752, 1753-1754, 1755-1756, 1757-1758, 1759-1760, 1761-1762, 1763-1764, 1765-1766, 1767-1768, 1769-1770, 1771-1772, 1773-1774, 1775-1776, 1777-1778, 1779-1780, 1781-1782, 1783-1784, 1785-1786, 1787-1788, 1789-1790, 1791-1792, 1793-1794, 1795-1796, 1797-1798, 1799-1800, 1801-1802, 1803-1804, 1805-1806, 1807-1808, 1809-1810, 1811-1812, 1813-1814, 1815-1816, 1817-1818, 1819-1820, 1821-1822, 1823-1824, 1825-1826, 1827-1828, 1829-1830, 1831-1832, 1833-1834, 1835-1836, 1837-1838, 1839-1840, 1841-1842, 1843-1844, 1845-1846, 1847-1848, 1849-1850, 1851-1852, 1853-1854, 1855-1856, 1857-1858, 1859-1860, 1861-1862, 1863-1864, 1865-1866, 1867-1868, 1869-1870, 1871-1872, 1873-1874, 1875-1876, 1877-1878, 1879-1880, 1881-1882, 1883-1884, 1885-1886, 1887-1888, 1889-1890, 1891-1892, 1893-1894, 1895-1896, 1897-1898, 1899-1900, 1901-1902, 1903-1904, 1905-1906, 1907-1908, 1909-1910, 1911-1912, 1913-1914, 1915-1916, 1917-1918, 1919-1920, 1921-1922, 1923-1924, 1925-1926, 1927-1928, 1929-1930, 1931-1932, 1933-1934, 1935-1936, 1937-1938, 1939-1940, 1941-1942, 1943-1944, 1945-1946, 1947-1948, 1949-1950, 1951-1952, 1953-1954, 1955-1956, 1957-1958, 1959-1960, 1961-1962, 1963-1964, 1965-1966, 1967-1968, 1969-1970, 1971-1972, 1973-1974, 1975-1976, 1977-1978, 1979-1980, 1981-1982, 1983-1984, 1985-1986, 1987-1988, 1989-1990, 1991-1992, 1993-1994, 1995-1996, 1997-1998, 1999-2000, 2001-2002, 2003-2004, 2005-2006, 2007-2008, 2009-2010, 2011-2012, 2013-2014, 2015-2016, 2017-2018, 2019-2020, 2021-2022, 2023-2024, 2025-2026, 2027-2028, 2029-2030, 2031-2032, 2033-2034, 2035-2036, 2037-2038, 2039-2040, 2041-2042, 2043-2044, 2045-2046, 2047-2048, 2049-2050, 2051-2052, 2053-2054, 2055-2056, 2057-2058, 2059-2060, 2061-2062, 2063-2064, 2065-2066, 2067-2068, 2069-2070, 2071-2072, 2073-2074, 2075-2076, 2077-2078, 2079-2080, 2081-2082, 2083-2084, 2085-2086, 2087-2088, 2089-2090, 2091-2092, 2093-2094, 2095-2096, 2097-2098, 209

Query Match 54.6%; Score 774.5; DB 1; Length 245;
Best Local Similarity 57.3%; Pred. No. 7.9e-57;
Matches 141; Conservative 32; Mismatches 72; Indels 1; Gaps 1;

QY 19 CGVPAITPALSINQRIYNGENAVPGSPWQVSLQDNTGFHFCGSLSPNNVYTAACHQV 78
DB 1 CGVPAIQVLSRLVINGEEAVPGSPWQVSLQDNTGFHFCGSLSPNNVYTAACHQV 60

QY 79 TPGRHFFVLGEYDRSSNAEPVQVLSIARATHPHNWNANTWNDLTLKLSAPARYTAQVS 138
DB 61 TTS-DVVYAGEFDQSSSERIQKIAKVFKNKSYNSLTINNDITLKLSTAAFSQSQTVS 119

QY 139 PVCLASTNEALPGLTCTVTTGWRISGVNVTAPRLOQVPLPLVTNNQCRQYNGARITDA 198
DB 120 AVCLPSASDFAGTCTVTTGWLTRYTNANTPDRLOQASLPPLSNTNCKRYWGTIKIDA 179

QY 199 MICAGSGGASCGDGGPLVCOKGNTWVLIGIVSWGTKNCNIQAPAMYTRVSKFSTWIN 258
DB 180 MICAGSGVSCMGDGGPLVCKKNGAWTLVGIVSWGSSSTCTSTPGVYARVATLVNMQ 239

QY 259 QVMAYN 264
DB 240 QTLAAN 245

RESULT 8
S72219
Chymotrypsin B - Atlantic cod (fragments)
C:Species: Gadus morhua (Atlantic cod)
C:Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 07-Aug-1998
C:Accession: S72219
R:Leth-Larsen, R.; Asgeirsson, B.; Thorolfsson, M.; Norregaard-Madsen, M.; Hojrup, P.
Biochim. Biophys. Acta 1297, 49-56, 1996
A:Title: Structure of chymotrypsin variant B from Atlantic cod, Gadus morhua.
A:Reference number: S72219; MUID:96439045; PMID:88411380
A:Accession: S72219
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-14;15-244 <LET>
C:Superfamily: trypsin; trypsin homology
F:15-237/Domain: trypsin homology <TRY>

Query Match 52.0%; Score 737; DB 2; Length 244;
Best Local Similarity 54.2%; Pred. No. 9.9e-54;
Matches 135; Conservative 43; Mismatches 63; Indels 8; Gaps 6;

QY 19 CGVPAITPALSINQRIYNGENAVPGSPWQVSLQDNTGFHFCGSLSPNNVYTAACHQV 78
DB 1 CGSPAIOQVT-GYAIYNGEEAVPHSPWQVSLQDNTGFHFCGSLSPNNVYTAACHQV 59

QY 79 TPGRHFFVLGEYDRSSNAEPVQVLSIARATHPHNWNANTWNDLTLKLSAPARYTAQVS 138
DB 60 RT-YHRVIVGEHDKASD-ENIQLKPSWVTHPKWDSRTINNDISLIKLSAPVGLTNVS 117

QY 139 PVCLASTNEALPGLTCTVTTGWRISGVNVTAPRLOQVPLPLVTNNQCRQYNGAR-I 195
DB 118 PVCLIGESSDFAPGKCVTSWGLTRYNAPG--TPNKLOQAALPLMSNECSQTWGNMI 175

QY 196 TDAMICAGSGGASCGDGGPLVCOKGNTWVLIGIVSWGTKNCNIQAPAMYTRVSKFST 255
DB 176 SDVMICAGAGATSCMGDGGPLVCOKDNTWTLVGIVSWGSSRCSTVTPYAYARVTELRG 235

QY 256 WINQVMAYN 264
DB 236 WVDQILAN 244

RESULT 9
A26823
pancreatic elastase II (EC 3.4.21.71) precursor - pig
N:Alternate names: pancrotopelptidase E
C:Species: Sus scrofa domestica (domestic pig)

C:Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 22-Jun-1999
C:Accession: A26823
R:Kawashima, I.; Tani, T.; Shimoda, K.; Takiguchi, Y.
DNA 6, 163-172, 1987
A:Title: Characterization of pancreatic elastase II cDNAs: two elastase II mRNAs are
A:Reference number: A90958; MUID:87217962; PMID:3646943
A:Accession: A26823
A:Molecule type: mRNA
A:Residues: 1-269 <KAW>
A:Cross-references: GB:M16651; NID:gl64441; PIDN:AAA1027.1; PID:gl644442
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-28/Domain: propeptide #status predicted <PRO>
F:29-269/Product: elastase II #status predicted <MAT>
F:29-262/Domain: trypsin homology <TRY>
F:73,121,216/Active site: His, Asp, Ser #status predicted

Query Match 37.9%; Score 537; DB 2; Length 269;
Best Local Similarity 43.1%; Pred. No. 4e-37;
Matches 118; Conservative 48; Mismatches 92; Indels 16; Gaps 11;

QY 2 LLSLSLTSLVLLGSSWCGVPAITPALSINQRIYNGENAVPGSPWQVSLQ-DNTG--FH 58
DB 1 MIRALLSTLVAG-ALSCGLPANLPQL---PRVVGEDARPNSWPQVSLQYDSSGQWRH 56

QY 59 FCGSLTSPNNVYTAACHQVTPGRHFVVLGEYDRSSNAEPVQVLSIARATHPHNWNANTM 118
DB 57 TCGTLVDQSWVLAHCHISSRTYRVLGRHSLSTNEPGSLAVKSVVHQMNSQL 116

QY 119 N--NDITLLKLKSPARYTAQVSPVCLASTNEALPGLTCTVTTGWRISGVNTPARLQ 176
DB 117 SGNNDIALLKLSPVSLTDRQLGCLPAAGTILPNKYCVYVTCWGRQLTNG-ASPDILQ 175

QY 177 VPLPLVTNNQCRQ--YNGARITDAMICAGSG-ASSCOGDSGGLVLCOKN-TWVLIGIV 232
DB 176 GQLLVVDYATCPKPGWGWSTVTKNICAGDGIISGDSGGLNCGGANGQMVHGIV 235

QY 233 SWGTK-NCN-IOAPAMYTRVSKFSTWINQVMAYN 264
DB 236 SFGSSLCGNYHKPSVTRVSNVIDWINSVIANN 269

RESULT 10
ELRT2
pancreatic elastase II (EC 3.4.21.71) precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 18-Aug-1982 #sequence_revision 18-Aug-1982 #text_change 24-Sep-1999
C:Accession: A00961
R:MacDonald, R.J.; Swift, G.H.; Quinto, C.; Swain, W.; Pictet, R.L.; Nikovits, W.; Ru
Biochemistry 21, 1453-1463, 1982
A:Title: Primary structure of two distinct rat pancreatic preproelastases determined
A:Reference number: A00960; MUID:82182967; PMID:6918221
A:Accession: A00961
A:Molecule type: mRNA
A:Residues: 1-271 <MAC>
A:Cross-references: GB:L00124; GB:J00731; NID:g204019; PIDN:AAA98780.1; PID:g204021
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; serine proteinase; zymogen
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-30/Domain: activation peptide #status predicted <ACT>
F:31-271/Product: elastase II #status predicted <MPT>
F:31-264/Domain: trypsin homology <TRY>
F:75,123,218/Active site: His, Asp, Ser #status predicted

Query Match 36.5%; Score 517; DB 1; Length 271;
Best Local Similarity 41.1%; Pred. No. 1.8e-35;
Matches 113; Conservative 46; Mismatches 100; Indels 16; Gaps 11;

QY 2 LLSLSLTSLVLLGSSWCGVPAITPALSINQRIYNGENAVPGSPWQVSLQ---DNTGF 57
DB 1 MIRTLLLSLVAG-ALSCGYP--TYEQVHDSRVVGGQEAQSPNSWPQVSLQYLSGSKWH 57

[illegible]

F:30-268/Product: pancreatic elastase isoform 1 #status predicted <MAT>
F:30-262/Domain: trypsin homology <TRY>

Query Match 34.98; Score 495.5; DB 2; Length 268;
Best Local Similarity 40.1%; Pred. No. 1.1e-33;
Matches 108; Conservative 46; Mismatches 102; Indels 13; Gaps 9;

Qy	3	LLSLTSLVLLGSSWGCGVPAITPALSYNORIVNGENAVPGSWPQVSLQ--DNTGPHF	59
Db	1	MLGITVLAALLACASSCGVPSPFNLS--ARVGGEDARPHSWPQISLQYLKNDTWKHT	58
Qy	60	CGSLISPNWYVTAACHCOVTPGRHFVVLGEYD-RSSNAEPVQVLSIARAI THPNMNANTM	118
Db	59	CGGTLIASNEFVLTAAHCISNRTYRVAVGKKNLEVEDEEGSLFVGVDTIHVHKKRNALLL	118
Qy	119	NNDLTLLKLSPARYTAQVSPVCLASTNEALPSGLTCVTTGCGRISGVGNVTPARLQOVV	178
Db	119	RNDIALIKLAEHVELSDTIQVACLPEKDSLPLKDYPCYVTGWRGLWTNGPIAD-KLQOGL	177
Qy	179	LPLVTVNOCRO--YVGARITDAMICAGGSA-SSCGDSGGPLVCOKGN-TWVLIGIVSW	234
Db	178	QPVDHATCSRIDWGWFFVKXTMVCAGGDGVISACNGDSGGPLNCQLENGSWEVFGIVSF	237
Qy	235	GT-KNCNI-QAPAMYTRVSKFSTWINQVM	261
Db	238	GSRRCNTRKKPVVYTVRSAYIDWINEKM	266

Search completed: December 20, 2002, 15:19:48
Job time : 17.6957 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:14:28 ; Search time 14.6087 Seconds
(without alignments)
1520.126 Million cell updates/sec

Title: US-09-856-319b-2_copy_1_231

Perfect score: 1221

Sequence: 1 MLLSLTLTLVLLGSSGGG.....GDSGGLVLCQKGTWVLIGI 231

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1214	99.4	264	2 I38136	chymotrypsin-like
2	714.5	58.5	263	2 A21195	chymotrypsin (EC 3.4.21.1)
3	701.5	57.5	263	2 S47537	chymotrypsin (EC 3.4.21.1)
4	689.5	56.5	263	2 A31299	chymotrypsin (EC 3.4.21.1)
5	670.5	54.9	263	1 KYRTB	chymotrypsin (EC 3.4.21.1)
6	666.5	54.6	245	1 KYBOB	chymotrypsin (EC 3.4.21.1)
7	657.5	53.8	245	1 KYBOA	chymotrypsin (EC 3.4.21.1)
8	653	53.5	244	2 S72219	chymotrypsin B - A
9	472	38.7	269	2 A6823	pancreatic elastase
10	453.5	37.1	270	2 B29934	pancreatic elastase
11	451	36.9	126	2 A23473	chymotrypsin-like
12	445	36.4	269	2 B26823	pancreatic elastase
13	438	35.9	271	1 ELRT2	pancreatic elastase
14	427	35.0	269	2 C26823	pancreatic elastase
15	420.5	34.4	270	2 A29934	pancreatic elastase
16	415	34.0	271	2 A25528	pancreatic elastase
17	410.5	33.6	343	1 A57014	proctasin (EC 3.4.21.1)
18	409	33.5	1524	2 T30337	polyprotein - Afri
19	404	33.1	460	2 B61545	plasmin (EC 3.4.21.1)
20	397	32.5	810	1 PLHU	plasmin (EC 3.4.21.1)
21	396	32.4	810	2 B30848	plasmin (EC 3.4.21.1)
22	395.5	32.4	1035	1 A43090	enteropeptidase (E
23	391	32.0	790	1 PLPG	plasmin (EC 3.4.21.1)
24	390.5	32.0	268	2 S68826	pancreatic elastase
25	390.5	32.0	268	2 S68825	pancreatic elastase
26	389.5	31.9	240	1 CPBOA3	procarboxypeptidase
27	389.5	31.9	812	1 PLMS	plasmin (EC 3.4.21.1)
28	389	31.9	455	2 A61545	plasmin (EC 3.4.21.1)
29	388	31.8	270	2 S56160	mast cell tryptase

30	386	31.6	786	1 A47547	serine proteinase
31	386	31.6	1004	2 T30338	oviductin (EC 3.4.21.1)
32	385	31.5	248	2 S55067	trypsin (EC 3.4.21.1)
33	384.5	31.5	238	2 S31779	trypsin (EC 3.4.21.1)
34	384.5	31.5	273	2 A47246	trypsin (EC 3.4.21.1)
35	383.5	31.4	246	1 TRRT1	trypsin (EC 3.4.21.1)
36	383.5	31.4	246	1 TRDGC	trypsin (EC 3.4.21.1)
37	382	31.3	1019	1 A56318	enteropeptidase (E
38	379	31.0	248	2 S55066	trypsin (EC 3.4.21.1)
39	378.5	30.9	638	1 KOHUP	plasma kallikrein
40	377.5	30.9	812	1 PLBO	plasmin (EC 3.4.21.1)
41	377.5	30.9	1113	2 JEO315	low-density lipopr
42	377	30.9	231	1 TRPCTR	trypsin (EC 3.4.21.1)
43	376.5	30.8	268	2 J01473	pancreatic elastase
44	376	30.8	266	1 ELRT1	pancreatic elastase
45	375.5	30.8	247	1 TRDG	trypsin (EC 3.4.21.1)

ALIGNMENTS

RESULT 1

I38136

chymotrypsin-like proteinase (EC 3.4.21.-) CTRL-1 - human

C:Species: Homo sapiens (man)

C>Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 22-Jun-1999

C:Accession: I38136

R:Larsen, F.; Solheim, J.; Kristensen, T.; Kolsto, A.B.; Prydz, H.

Hum. Mol. Genet. 2, 1589-1595, 1993

A:Title: A tight cluster of five unrelated human genes on chromosome 16q22.1.

A:Reference number: I38135; MUID:94093544; PMID:8268911

A:Accession: I38136

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Residues: 1-264 <RES>

A:Cross-references: EMBL:X71874; NID:g406226; PIDN:CAA50710.1; PID:g406228

C:Genetics:

A:Gene: GDB:CTRL

A:Cross-references: GDB:204061

A:Map position: 16q22.1-16q22.1

A:Introns: 18/1; 52/3; 79/2; 106/3; 167/1; 211/3

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; serine proteinase

F:34-257/Domain: trypsin homology <TRY>

F:75.121.214/Active site: His, Asp, Ser #status predicted

Query Match

Best Local Similarity 99.4%; Score 1214; DB 2; Length 264;

Matches 230; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLLSLTLTLVLLGSSGGGIPAIKPAISQRIVNGENAVLGSWPQVSLQDSSGFHFC 60

Db 1 MLLSLTLTLVLLGSSGGGIPAIKPAISQRIVNGENAVLGSWPQVSLQDSSGFHFC 60

QY 61 GGLISQSWVTAACHNCVSPGRHFVVLGYDRSSNAEPLQVLSVRAITHPSNNTMNN 120

Db 61 GGLISQSWVTAACHNCVSPGRHFVVLGYDRSSNAEPLQVLSVRAITHPSNNTMNN 120

QY 121 DVTLLKLSPAQYTTTRISPVCLASNEALTEGLTCVTGNGRLSGVGNVTPAHLQVALP 180

Db 121 DVTLLKLSPAQYTTTRISPVCLASNEALTEGLTCVTGNGRLSGVGNVTPAHLQVALP 180

QY 181 LVTVNOCROYWSSITDSMICAGGAGSCGDSGGGLVLCQKGTWVLIGI 231

Db 181 LVTVNOCROYWSSITDSMICAGGAGSCGDSGGGLVLCQKGTWVLIGI 231

RESULT 2

A21195

chymotrypsin (EC 3.4.21.1) 2 precursor - dog

C:Species: Canis lupus familiaris (dog)

C>Date: 27-Jul-1990 #sequence_revision 27-Jul-1990 #text_change 22-Jun-1999

C:Accession: A21195

R:Pinsky, S.D.; LaForge, K.S.; Luc, V.; Scheele, G.

Proc. Natl. Acad. Sci. U.S.A. 80, 7486-7490, 1983

A:Title: Identification of cDNA clones encoding secretory isoenzyme forms: sequence data

A:Reference number: A21195; MUID:84170253; PMID:6584866

A:Accession: A21195

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-263 <PIN>

A:Cross-references: GB:K01173; NID:g163945; PIDN:AAA30841.1; PID:g163946

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; protein digestion; serine proteinase

F:34-256/Domain: trypsin homology <TRY>

F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 58.5%; Score 714.5; DB 2; Length 263;

Best Local Similarity 56.7%; Pred. No. 5e-53;

Matches 131; Conservative 35; Mismatches 64; Indels 1; Gaps 1;

QY 1 MLLSLTSLVLGSGGCGIPALPALSQRIVNGENAVLGSWPQVSLQDSSGFHFC 60

Db 1 MAFLLWLLSCFALLGTAFGCGVPAIQPVLSGLSRVNGEDAVPGSWPQVSLQDSTGFHFC 60

QY 61 GGLISQSVVVTAAHCNVSPGRHFVYLGEYDRSSNAEPLOVLSVRAITHPSWNSTMMN 120

Db 61 GGLISQSVVVTAAHCNVSPGRHFVYLGEYDRSSNAEPLOVLSVRAITHPSWNSTMMN 119

QY 121 DVTLKLLASPAQVYTRISPVCLASSNEALTEGLTCVTTGWLGRSGVGNVTPAHLQVVALP 180

Db 120 DYTLLKLLATPARESKVSAVCLPQATDDFPAGTLCVTTGWLGRKHNANTPDKLQQAALP 179

QY 181 LVTVNOCROYWDSITDSMTICAGAGASSCGDGGPLVCQKGNWTVLIGI 231

Db 180 LLSNAECKKFWGSKITDLMVCAGASGVSCMGDSGGPLVCQKGNWTVLIGI 230

RESULT 3

S47537

chymotrypsin (EC 3.4.21.1) precursor - Atlantic cod

C:Species: Gadus morhua (Atlantic cod)

C:Date: 26-Dec-1994 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999

C:Accession: S47537; S43163

R:Gudmundsdottir, A.; Oskarsson, S.; Eakin, A.E.; Craik, C.S.; Bjarnason, J.B.

Biochim. Biophys. Acta 1219, 211-214, 1994

A:Title: Atlantic cod cDNA encoding a psychrophilic chymotrypsinogen.

A:Reference number: S47537; MUID:94368860; PMID:8086467

A:Accession: S47537

A:Molecule type: mRNA

A:Residues: 1-263 <GUD>

A:Cross-references: EMBL:X78490; NID:g468750; PIDN:CAA55242.1; PID:g468751

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; protein digestion; serine proteinase

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-263/Product: chymotrypsin #status predicted <PRO>

F:34-256/Domain: trypsin homology <TRY>

F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 57.5%; Score 701.5; DB 2; Length 263;

Best Local Similarity 58.6%; Pred. No. 6.3e-52;

Matches 129; Conservative 33; Mismatches 57; Indels 1; Gaps 1;

QY 12 LLGSSGCGIPALPALSQRIVNGENAVLGSWPQVSLQDSSGFHFCGGLISQSVV 71

Db 12 LFRITGCGRPALSPVITGVSRVNGEAVPHSMQVSLQDQTFHFCGGLINENVV 71

QY 72 TAAHCNVSPGRHFVYLGEYDRSSNAEPLOVLSVRAITHPSWNSTMMNDVTLKLLASPA 131

Db 72 TAAHCNVK-NYHRAVLGEHDSRNSGQVMTGVQFKHPRYNGFTINNDILLVKLATPA 130

QY 132 QYTRISPVCLASSNEALTEGLTCVTTGWLGRSGVGNVTPAHLQVVALPVTNOCROYW 191

Db 131 TLNMRVSPVCLATDDTFEGGCMKCVTSGLTTRYNAADTPALLQQAALPLLTNEQCKKF 190

QY 192 DSSITDSMTICAGAGASSCGDGGPLVCQKGNWTVLIGI 231

Db 191 GNKISDLMICAGAGASSCMGDSGGPLVCQKAGSWTLVGI 230

RESULT 4

A31299

chymotrypsin (EC 3.4.21.1) precursor - human

C:Species: Homo sapiens (man)

C:Date: 08-Jun-1989 #sequence_revision 08-Jun-1989 #text_change 22-Jun-1999

C:Accession: A31299

R:Tomita, N.; Izumoto, Y.; Horii, A.; Doi, S.; Yokouchi, H.; Ogawa, M.; Mori, T.; Mat

Biochem. Biophys. Res. Commun. 158, 569-575, 1989

A:Title: Molecular cloning and nucleotide sequence of human pancreatic prechymotrypsin

A:Reference number: A31299; MUID:89134264; PMID:2917002

A:Accession: A31299

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-263 <TOM>

A:Cross-references: GB:M24400; NID:g181189; PIDN:AAA52128.1; PID:g181190

C:Genetics:

A:Gene: GDB:CTRB1; CTRB

A:Cross-references: GDB:119820; OMIM:118890

A:Map position: 16q23.1-16q23.1

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; protein digestion; serine proteinase

F:34-256/Domain: trypsin homology <TRY>

F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 56.5%; Score 689.5; DB 2; Length 263;

Best Local Similarity 55.4%; Pred. No. 6.5e-51;

Matches 128; Conservative 35; Mismatches 67; Indels 1; Gaps 1;

QY 1 MLLSLTSLVLGSGGCGIPALPALSQRIVNGENAVLGSWPQVSLQDSSGFHFC 60

Db 1 MAFLLWLLSCFALLGTAFGCGVPAIQPVLSGLSRVNGEDAVPGSWPQVSLQDSTGFHFC 60

QY 61 GGLISQSVVVTAAHCNVSPGRHFVYLGEYDRSSNAEPLOVLSVRAITHPSWNSTMMN 120

Db 61 GGLISQSVVVTAAHCNVSPGRHFVYLGEYDRSSNAEPLOVLSVRAITHPSWNSTMMN 119

QY 121 DVTLKLLASPAQVYTRISPVCLASSNEALTEGLTCVTTGWLGRSGVGNVTPAHLQVVALP 180

Db 120 DYTLLKLLATPARESKVSAVCLPQATDDFPAGTLCVTTGWLGRKHNANTPDKLQQAALP 179

QY 181 LVTVNOCROYWDSITDSMTICAGAGASSCGDGGPLVCQKGNWTVLIGI 231

Db 180 LLSNAECKKSWGRITDLMVCAGASGVSCMGDSGGPLVCQKGNWTVLIGI 230

RESULT 5

KYRTB

chymotrypsin (EC 3.4.21.1) B precursor - rat

N:Alternate names: chymotrypsinogen B

C:Species: Rattus norvegicus (Norway rat)

C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 18-Jun-1999

C:Accession: A22658

R:Bell, G.I.; Quinto, C.; Quiroga, M.; Valenzuela, P.; Craik, C.S.; Rutter, W.J.

J. Biol. Chem. 259, 14265-14270, 1984

A:Title: Isolation and sequence of a rat chymotrypsin B gene.

A:Reference number: A22658; MUID:85054881; PMID:6209274

A:Accession: A22658

A:Molecule type: DNA

A:Residues: 1-263 <BEL>

A:Cross-references: GB:K02298; NID:g203653; PIDN:AAA98732.1; PID:g203654

C:Genetics:

A:Introns: 18/1; 52/3; 79/2; 105/3; 166/1; 210/3

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-33/Domain: propeptide #status predicted <PRO>

F:34-263/Product: chymotrypsin B #status predicted <MAT>

F:75,120,213/Active site: trypsin homology <TRY>

F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 54.98; Score 670.5; DB 1; Length 263;
Best Local Similarity 53.28; Pred. No. 2.6e-49;
Matches 123; Conservative 35; Mismatches 72; Indels 1; Gaps 1;

Qy 1 MLLSLTSLVLLSGMCGGIPALPAKSFQRIVNGENAVLGSWPQVSLQDSSGFHFC 60
Db 1 MAFLWLVSFCALVATGCGVPTIOPVLTLGLSRVNGEDAIPGSWPQVSLQDKTGFHFC 60

Qy 61 GGSLSIOSWVYTAHNCVNSPGRHFVVLGEYDRSSNAEPLQVLSVRAITHPSWNTMNN 120
Db 61 GGSLSIEDWVYTAHNCVNSPGRHFVVLGEYDRSSNAEPLQVLSVRAITHPSWNTMNN 119

Qy 121 DVTLLKLASPAQVYTRTSIPVCLASSNEALTEGLTCVTTGWLRSVGNVTFAHLQVVALP 180
Db 120 DITLLKLATPAQFSETVSAVCLPNVDPPETVCATITGKTKYNALKTEPEKIQQAALP 179

Qy 181 LVTYNOCROYWDSSTDMTCAGAGASSCGDGGPLVCQKGNWTWVLI 231
Db 180 IVSEADCKKSGSKITDVTMTAGASGVSSCMGDSGGPLVCQKGNWTWVLI 230

RESULT 6

KYBOB

chymotrypsin (EC 3.4.21.1) B precursor - bovine
N:Alternate names: chymotrypsinogen B
C:Species: Bos primigenius taurus (cattle)
C:Date: 08-Oct-1981 #sequence_revision 08-Oct-1981 #text_change 18-Jul-1997
C:Accession: A00953
Nature 218, 343-346, 1968
A:Title: Structure of chymotrypsinogen B compared with chymotrypsinogen A and trypsinogen A
A:Reference number: A00953; MUID:68238908; PMID:5649671
A:Accession: A00953
A:Molecule type: protein
A:Residues: 1-245 <SMI>
C:Comment: Chymotrypsinogen B is synthesized, along with chymotrypsinogen A, in the acinar cells of the pancreas.
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F:1-15/Domain: propeptide #status experimental <PRO>
F:16-245/Product: chymotrypsin B #status experimental <MAT>
F:16-238/Domain: trypsin homology <TRY>
F:1-122,42-58,136-201,168-182,191-220/Disulfide bonds: #status experimental
F:57,102,195/Active site: His, Asp, Ser #status experimental

Query Match 54.68; Score 666.5; DB 1; Length 245;
Best Local Similarity 57.38; Pred. No. 5.2e-49;
Matches 122; Conservative 28; Mismatches 62; Indels 1; Gaps 1;

Qy 19 CGIPALPAKSFQRIVNGENAVLGSWPQVSLQDSSGFHFCGSLISOSWVYTAHNCV 78
Db 1 CGVPAIQVLSGLARVNGEDAVPGSWPQVSLQDSTGFHFCGSLISEDWVYTAHNCV 60

Qy 79 SPGRHFVVLGEYDRSSNAEPLQVLSVRAITHPSWNTMNNVTLKLASPAQVYTRIS 138
Db 61 TTS-DVVYAGEFDQGLETDQVLKIGKVFKNPFSILTVRNDITLLKLATPAQFSEVTS 119

Qy 139 PVCLASSNEALTEGLTCVTTGWLRSVGNVTFAHLQVVALPVTYNOCROYWDSSTIDS 198
Db 120 AVCLPSADEDFPAGMLCATTCGKTKYNALKTPKLOQAATLPVSNITDCRYWGSRTVDV 179

Qy 199 MICAGGAGASSCGDGGPLVCQKGNWTWVLI 231
Db 180 MICAGASVSSCMGDSGGPLVCQKGNWTWVLI 212

RESULT 7

KYBOA

chymotrypsin (EC 3.4.21.1) A precursor - bovine
N:Alternate names: chymotrypsinogen A
C:Species: Bos primigenius taurus (cattle)
C:Date: 07-May-1981 #sequence_revision 07-May-1981 #text_change 07-May-1999

C:Accession: A90235; A93158; S29650; A00952
R:Brown, J.R.; Hartley, B.S.
Biochem. J. 101, 214-228, 1966
A:Title: Location of disulphide bridges by diagonal paper electrophoresis. The disulphide bridges are located at positions 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993, 994, 995, 996, 997, 998, 999, 1000, 1001, 1002, 1003, 1004, 1005, 1006, 1007, 1008, 1009, 1010, 1011, 1012, 1013, 1014, 1015, 1016, 1017, 1018, 1019, 1020, 1021, 1022, 1023, 1024, 1025, 1026, 1027, 1028, 1029, 1030, 1031, 1032, 1033, 1034, 1035, 1036, 1037, 1038, 1039, 1040, 1041, 1042, 1043, 1044, 1045, 1046, 1047, 1048, 1049, 1050, 1051, 1052, 1053, 1054, 1055, 1056, 1057, 1058, 1059, 1060, 1061, 1062, 1063, 1064, 1065, 1066, 1067, 1068, 1069, 1070, 1071, 1072, 1073, 1074, 1075, 1076, 1077, 1078, 1079, 1080, 1081, 1082, 1083, 1084, 1085, 1086, 1087, 1088, 1089, 1090, 1091, 1092, 1093, 1094, 1095, 1096, 1097, 1098, 1099, 1100, 1101, 1102, 1103, 1104, 1105, 1106, 1107, 1108, 1109, 1110, 1111, 1112, 1113, 1114, 1115, 1116, 1117, 1118, 1119, 1120, 1121, 1122, 1123, 1124, 1125, 1126, 1127, 1128, 1129, 1130, 1131, 1132, 1133, 1134, 1135, 1136, 1137, 1138, 1139, 1140, 1141, 1142, 1143, 1144, 1145, 1146, 1147, 1148, 1149, 1150, 1151, 1152, 1153, 1154, 1155, 1156, 1157, 1158, 1159, 1160, 1161, 1162, 1163, 1164, 1165, 1166, 1167, 1168, 1169, 1170, 1171, 1172, 1173, 1174, 1175, 1176, 1177, 1178, 1179, 1180, 1181, 1182, 1183, 1184, 1185, 1186, 1187, 1188, 1189, 1190, 1191, 1192, 1193, 1194, 1195, 1196, 1197, 1198, 1199, 1200, 1201, 1202, 1203, 1204, 1205, 1206, 1207, 1208, 1209, 1210, 1211, 1212, 1213, 1214, 1215, 1216, 1217, 1218, 1219, 1220, 1221, 1222, 1223, 1224, 1225, 1226, 1227, 1228, 1229, 1230, 1231, 1232, 1233, 1234, 1235, 1236, 1237, 1238, 1239, 1240, 1241, 1242, 1243, 1244, 1245, 1246, 1247, 1248, 1249, 1250, 1251, 1252, 1253, 1254, 1255, 1256, 1257, 1258, 1259, 1260, 1261, 1262, 1263, 1264, 1265, 1266, 1267, 1268, 1269, 1270, 1271, 1272, 1273, 1274, 1275, 1276, 1277, 1278, 1279, 1280, 1281, 1282, 1283, 1284, 1285, 1286, 1287, 1288, 1289, 1290, 1291, 1292, 1293, 1294, 1295, 1296, 1297, 1298, 1299, 1300, 1301, 1302, 1303, 1304, 1305, 1306, 1307, 1308, 1309, 1310, 1311, 1312, 1313, 1314, 1315, 1316, 1317, 1318, 1319, 1320, 1321, 1322, 1323, 1324, 1325, 1326, 1327, 1328, 1329, 1330, 1331, 1332, 1333, 1334, 1335, 1336, 1337, 1338, 1339, 1340, 1341, 1342, 1343, 1344, 1345, 1346, 1347, 1348, 1349, 1350, 1351, 1352, 1353, 1354, 1355, 1356, 1357, 1358, 1359, 1360, 1361, 1362, 1363, 1364, 1365, 1366, 1367, 1368, 1369, 1370, 1371, 1372, 1373, 1374, 1375, 1376, 1377, 1378, 1379, 1380, 1381, 1382, 1383, 1384, 1385, 1386, 1387, 1388, 1389, 1390, 1391, 1392, 1393, 1394, 1395, 1396, 1397, 1398, 1399, 1400, 1401, 1402, 1403, 1404, 1405, 1406, 1407, 1408, 1409, 1410, 1411, 1412, 1413, 1414, 1415, 1416, 1417, 1418, 1419, 1420, 1421, 1422, 1423, 1424, 1425, 1426, 1427, 1428, 1429, 1430, 1431, 1432, 1433, 1434, 1435, 1436, 1437, 1438, 1439, 1440, 1441, 1442, 1443, 1444, 1445, 1446, 1447, 1448, 1449, 1450, 1451, 1452, 1453, 1454, 1455, 1456, 1457, 1458, 1459, 1460, 1461, 1462, 1463, 1464, 1465, 1466, 1467, 1468, 1469, 1470, 1471, 1472, 1473, 1474, 1475, 1476, 1477, 1478, 1479, 1480, 1481, 1482, 1483, 1484, 1485, 1486, 1487, 1488, 1489, 1490, 1491, 1492, 1493, 1494, 1495, 1496, 1497, 1498, 1499, 1500, 1501, 1502, 1503, 1504, 1505, 1506, 1507, 1508, 1509, 1510, 1511, 1512, 1513, 1514, 1515, 1516, 1517, 1518, 1519, 1520, 1521, 1522, 1523, 1524, 1525, 1526, 1527, 1528, 1529, 1530, 1531, 1532, 1533, 1534, 1535, 1536, 1537, 1538, 1539, 1540, 1541, 1542, 1543, 1544, 1545, 1546, 1547, 1548, 1549, 1550, 1551, 1552, 1553, 1554, 1555, 1556, 1557, 1558, 1559, 1560, 1561, 1562, 1563, 1564, 1565, 1566, 1567, 1568, 1569, 1570, 1571, 1572, 1573, 1574, 1575, 1576, 1577, 1578, 1579, 1580, 1581, 1582, 1583, 1584, 1585, 1586, 1587, 1588, 1589, 1590, 1591, 1592, 1593, 1594, 1595, 1596, 1597, 1598, 1599, 1600, 1601, 1602, 1603, 1604, 1605, 1606, 1607, 1608, 1609, 1610, 1611, 1612, 1613, 1614, 1615, 1616, 1617, 1618, 1619, 1620, 1621, 1622, 1623, 1624, 1625, 1626, 1627, 1628, 1629, 1630, 1631, 1632, 1633, 1634, 1635, 1636, 1637, 1638, 1639, 1640, 1641, 1642, 1643, 1644, 1645, 1646, 1647, 1648, 1649, 1650, 1651, 1652, 1653, 1654, 1655, 1656, 1657, 1658, 1659, 1660, 1661, 1662, 1663, 1664, 1665, 1666, 1667, 1668, 1669, 1670, 1671, 1672, 1673, 1674, 1675, 1676, 1677, 1678, 1679, 1680, 1681, 1682, 1683, 1684, 1685, 1686, 1687, 1688, 1689, 1690, 1691, 1692, 1693, 1694, 1695, 1696, 1697, 1698, 1699, 1700, 1701, 1702, 1703, 1704, 1705, 1706, 1707, 1708, 1709, 1710, 1711, 1712, 1713, 1714, 1715, 1716, 1717, 1718, 1719, 1720, 1721, 1722, 1723, 1724, 1725, 1726, 1727, 1728, 1729, 1730, 1731, 1732, 1733, 1734, 1735, 1736, 1737, 1738, 1739, 1740, 1741, 1742, 1743, 1744, 1745, 1746, 1747, 1748, 1749, 1750, 1751, 1752, 1753, 1754, 1755, 1756, 1757, 1758, 1759, 1760, 1761, 1762, 1763, 1764, 1765, 1766, 1767, 1768, 1769, 1770, 1771, 1772, 1773, 1774, 1775, 1776, 1777, 1778, 1779, 1780, 1781, 1782, 1783, 1784, 1785, 1786, 1787, 1788, 1789, 1790, 1791, 1792, 1793, 1794, 1795, 1796, 1797, 1798, 1799, 1800, 1801, 1802, 1803, 1804, 1805, 1806, 1807, 1808, 1809, 1810, 1811, 1812, 1813, 1814, 1815, 1816, 1817, 1818, 1819, 1820, 1821, 1822, 1823, 1824, 1825, 1826, 1827, 1828, 1829, 1830, 1831, 1832, 1833, 1834, 1835, 1836, 1837, 1838, 1839, 1840, 1841, 1842, 1843, 1844, 1845, 1846, 1847, 1848, 1849, 1850, 1851, 1852, 1853, 1854, 1855, 1856, 1857, 1858, 1859, 1860, 1861, 1862, 1863, 1864, 1865, 1866, 1867, 1868, 1869, 1870, 1871, 1872, 1873, 1874, 1875, 1876, 1877, 1878, 1879, 1880, 1881, 1882, 1883, 1884, 1885, 1886, 1887, 1888, 1889, 1890, 1891, 1892, 1893, 1894, 1895, 1896, 1897, 1898, 1899, 1900, 1901, 1902, 1903, 1904, 1905, 1906, 1907, 1908, 1909, 1910, 1911, 1912, 1913, 1914, 1915, 1916, 1917, 1918, 1919, 1920, 1921, 1922, 1923, 1924, 1925, 1926, 1927, 1928, 1929, 1930, 1931, 1932, 1933, 1934, 1935, 1936, 1937, 1938, 1939, 1940, 1941, 1942, 1943, 1944, 1945, 1946, 1947, 1948, 1949, 1950, 1951, 1952, 1953, 1954, 1955, 1956, 1957, 1958, 1959, 1960, 1961, 1962, 1963, 1964, 1965, 1966, 1967, 1968, 1969, 1970, 1971, 1972, 1973, 1974, 1975, 1976, 1977, 1978, 1979, 1980, 1981, 1982, 1983, 1984, 1985, 1986, 1987, 1988, 1989, 1990, 1991, 1992, 1993, 1994, 1995,


```
Query Match 37.1%; Score 453.5; DB 2; Length 270;
Best Local Similarity 42.5%; Pred. No. 5.8e-31;
Matches 102; Conservative 36; Mismatches 87; Indels 15; Gaps 8;

Qy 2 LLSLTSLVLGSSWGGGIPAIKPAKLSFSQRIYNGENAVLGSPWQVSLQ-DSSG---FH 58
Db 1 MMLRLSSLLLVAVASGYGPPSSRP-----SSRVNGEDAVPYSWPQVSLQYKSGSPYH 56

Qy 59 FCGSLISQSWVVAACNCVSPGRHFVVLGEYDRSSNAEPLOVLSVRA---ITHPSWNST 116
Db 57 TCGSLIADPWVVTAGHCISSTRYQVVLGEYDRVKEGPEQVPIPIGSGDLFVPLMNR 116

Qy 117 TM--NNDTLLKSPAQYTRISPVCLASSNEALTEGLTCVTTGWRLSGVGNVTPAHL 174
Db 117 CVACGNDIALIKLSRAQLGDAVOLASLPAGDILNPETCYITGWGRLYTNGPL-PPKL 175

Qy 175 QOVALPLVTNNOCRO--YWDSSITDSMTICAGAGASSCGQSDGGLPLVC-QKGNVTWLIGI 231
Db 176 QEALLPVVDYEHCSRWNWNGSVKRTMVCAGDIRSGCNGSDGGLPLNCPTEDGGVQVHGV 235

RESULT 11
A23473
Chymotrypsin-like proteinase (EC 3.4.21.-) - pig (tentative sequence) (fragments)
N:Alternate names: pancreatic elastase II [misidentification]
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 31-Mar-2000
C:Accession: A23473
R:Vered, M.; Gertler, A.; Burstein, Y.
Int. J. Pept. Protein Res. 27, 183-190, 1986
A:Reference number: A23473; MUID:86194934; PMID:3634756
A:Accession: A23473
A:Molecule type: protein
A:Residues: 1-126 <VER>
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase

Query Match 36.9%; Score 451; DB 2; Length 126;
Best Local Similarity 43.7%; Pred. No. 3.9e-31;
Matches 93; Conservative 13; Mismatches 11; Indels 96; Gaps 3;

Qy 19 CGIPAIKPAKLSFSQRIYNGENAVLGSPWQVSLQDSSGFHFCGSLISQSWVVAACNCV 78
Db 1 CGVPAIKPALB-----IVBG2BAVPCSWPZWVSL2BSBGFHFCGSLISZ----- 45

Qy 79 SPGRHFVVLGEYDRSSNAEPLOVLSVRAITHPSWNSTMTMNDVTLKLSAPQYTRIS 138
Db 46 -----MNDLTLLKLSAPQYTRIG 66

Qy 139 PVCLASSNEALTEGLTCVTTGWRLSGVGNVTPAHLQOVALPLVTNNOCROYWDSSTIDS 198
Db 67 PVCLASGEALPEXLPVCV-----MNDLTLLKLSAPQYTRIG 66

Qy 199 MICAGGASSCGQSDGGLPLVCQKGNVTWLIGI 231
Db 85 MICAGASSCGQSDGGLPLVCXKGDXTWLXGI 117

RESULT 12
B26823
pancreatic elastase II (EC 3.4.21.71) A precursor - human
C:Species: Homo sapiens (man)
C>Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 20-Jun-2000
C:Accession: B26823; A27432; A41431; S34491
R:Kawashima, I.; Tani, T.; Shimoda, K.; Takiguchi, Y.
DNA 6, 163-172, 1987
A:Title: Characterization of pancreatic elastase II cDNAs: two elastase II mRNAs are expressed
A:Reference number: A90958; MUID:87217962; PMID:3646943
A:Accession: B26823
A:Molecule type: mRNA
A:Residues: 1-269 <KAW>
A:Cross-references: GB:M16652; NID:g182057; PIDN:AAA52380.1; PID:g182058
```

```
R:Fletcher, T.S.; Shen, W.F.; Largman, C.
Biochemistry 26, 7256-7261, 1987
A:Title: Primary structure of human pancreatic elastase 2 determined by sequence analysis
A:Reference number: A27432; MUID:88107669; PMID:3427074
A:Accession: A27432
A:Molecule type: mRNA
A:Residues: 1-269 <FLE>
A:Cross-references: GB:M16631; NID:g182022; PIDN:AAA52374.1; PID:g182023
R:Shirasu, Y.; Yoshida, H.; Matsuki, S.; Takemura, K.; Ikeda, N.; Shimada, Y.; Ozawa, J.
Biochem. 102, 1555-1563, 1987
A:Title: Molecular cloning and expression in Escherichia coli of a cDNA encoding human pancreatic elastase
A:Reference number: A41431; MUID:88198076; PMID:2834346
A:Accession: A41431
A:Molecule type: mRNA
A:Residues: 1-201 'V', 203-269 <SHI>
A:Cross-references: GB:D00236; NID:g219619; PIDN:BAA00165.1; PID:g219620
A:Note: The authors translated the codon GTG for residue 202 as Cys
R:Moulard, M.; Michon, T.; Kerfelec, B.; Chapus, C.
FEBS Lett. 261, 179-183, 1990
A:Title: Further studies on the human pancreatic binary complexes involving procarboxypeptidase
A:Reference number: S08253; MUID:90169111; PMID:2307232
A:Accession: S34491
A:Molecule type: protein
A:Residues: 'X', 18-50 <MOU>
C:Genetics:
A:Gene: GDB:ELAI
A:Cross-references: GDB:I19866; OMIM:130120
A:Map position: 12pter-12qter
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; serine proteinase
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-28/Domain: propeptide #status predicted <PRO>
F:29-269/Product: pancreatic elastase IIA #status predicted <NAT>
F:29-262/Domain: trypsin homology <TRY>
F:73,121,216/Active site: His, Asp, Ser #status predicted

Query Match 36.4%; Score 445; DB 2; Length 269;
Best Local Similarity 43.9%; Pred. No. 3e-30;
Matches 105; Conservative 28; Mismatches 92; Indels 14; Gaps 8;

Qy 2 LLSLTSLVLGSSWGGGIPAIKPAKLSFSQRIYNGENAVLGSPWQVSLQDSSG---FH 58
Db 1 MIRTLLSTLVAG-ALSCGDPTYP---YVTRVVGGEARPNSPWQVSLQYSSNGKWKYH 56

Qy 59 FCGSLISQSWVVAACNCVSPGRHFVVLGEYDRSSNAEPLOVLSVRAITHPSWNSTMT 118
Db 57 TCGSLIANSVLTAACHCISSTRYRVLGRHNLVVAESGLAVSVSVIHKDWSNOI 116

Qy 119 N--NDVTLLKLSAPQYTRISPVCLASSNEALTEGLTCVTTGWRLSGVGNVTPAHLQO 176
Db 117 SKGNDIALKLANPVSITDKIQACLPAGTILPNNYPCYVTGWRLQTNGAV-PDV LQO 175

Qy 177 VALPLVTNNOCRO---YWDSSITDSMICAGGAGA--SSCGSDGGLPLVCQKGN-TWVLIGI 231
Db 176 GRLLVVDYATCGSSAWWSSVSKTSMICAGGCVISSCNGSGGLPLNCQASDGRWQVHGI 234

RESULT 13
ELRT2
pancreatic elastase II (EC 3.4.21.71) precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 18-Aug-1982 #sequence_revision 18-Aug-1982 #text_change 24-Sep-1999
C:Accession: A00961
R:Macdonald, R.J.; Swift, G.H.; Quinto, C.; Swain, W.; Pictet, R.L.; Nikovits, W.; R.
Biochemistry 21, 1453-1463, 1982
A:Title: Primary structure of two distinct rat pancreatic preproelastases determined
A:Reference number: A00960; MUID:82182967; PMID:6918221
A:Accession: A00961
A:Molecule type: mRNA
A:Residues: 1-271 <MAC>
A:Cross-references: GB:L00124; GB:J00731; NID:g204019; PIDN:AAA98780.1; PID:g204021
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; serine proteinase; zymogen
```

F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-30/Domain: activation peptide #status predicted <APT>
F;31-271/Product: elastase II #status predicted <NPT>
F;31-264/Domain: trypsin homology <TR>
F;75,123,218/Active site: His, Asp, Ser #status predicted

Query Match 35.9%; Score 438; DB 1; Length 271;
Best Local Similarity 41.8%; Pred. No. 1.2e-29;

Qy	2	LLLSLTLVLSSGSCGIPATKPAISPSORIVNENAVLSWPQVSLQD - SSG - FH 58
Db	1	MIRTLLSALVAG - ALSGCGPTTVEVQHDVS - RYVGQGEASPSPWPQVSLQVSSGKWH 58
Qy	59	FCGSLISQSVMVTAACHNVSPCRHFVLGEYDRSSNAEPQLVLSYSRAITHPSNWNSTM 118
Db	59	TCGSLVANNWVLTAAHCTNSSTYRVILGRHSLSTSESGSLAVQSKVLVVEKWNNAQL 118
Qy	119	N - -NDVTLKLLASPAQVTTTRISPVCLASSNEALTEGLTCTVTGWRGLSGVGNVTPAHLQ 176
Db	119	SNGNDIALVKLASPVALTSKIQTACLPPACTILPNVPCYVTGWRGLQTNG - ATPDVLQ 177
Qy	177	VALPLVTVNQC - -RQWDDSSITDSMTACGAG - ASSCQDGSGLPVCQKGN - TWVLIGI 231
Db	178	GRLLVVDYATCTSSASMMGSSVKTNMYCAGDGYTSCNGDGSGLPNCOA SNGOWVHGI 236

RESULT 14

pancreatic elastase II (EC 3.4.21.71) B precursor - human
C26823
N;Alternate names: pancratopeptidase E
C;Species: Homo sapiens (man)
C;Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 22-Jun-1999
C;Accession: C26823
R;Kawashima, I.; Tani, T.; Shimoda, K.; Takiauchi, Y.

Query Match 35.0%; Score 427; DB 2; Length 269;
Best Local Similarity 41.0%; Pred. No. 1e-28;
Matches 98: Conservative 36; Mismatches 91; Indels 14; Gaps 8;

Qy	2	LLLSLTUSLVLLSSGCGGTPATKPALESFSQRTVNGENAVLGSWPQVSLQDSSG---PH 58
Db	1	MIRPFLLSLTVAG-ALSCGVSYTAPDMS---RMLGGEERAPNSWPQVSLQYSSNGQWTH 56
Qy	59	FCGSLISQSNVYTAACHNCPGRHFHVLGEXDRSSNAEPLOVLVSYSRAITHPSNNSTIM 118
Db	57	TCGSLTANSWLTAAHCISSSRIYRMVLQGHLYVAESGSLAVSVKILVHVHDMNSOV 116
Qy	119	N--NDVTLKLASPAQYTTIRISPVCLASSNEALTEGLTCVTTGWRLSGVGNVTPAHQQ 176
Db	117	SKGNDIALLKANPVSLTDKLIQALCPAGTILPNVYPCVVTGWGRLOING-ALPDLAQ 175
Qy	177	VALPLVTVMOCRO--YWDSSITDSMTICAGAGA-SSCGDSDSGPLVCQKGN-TWVLIGI 231
Db	176	GRLLVDPYATYCSSGNGWSTVKNIMICAGDGYICITCNGDSDGPLNCAQASDGRWEYHGI 234

RESULT 15

A29934

pancreatic elastase (EC 3.4.21.36) IIIA precursor - human
N:Alternate names: protease E
C:Species: Homo sapiens (man)
C:Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 20-Jun-2000
C:Accession: A29934; JX0045
R:Tani, T.; Ohsumi, J.; Mita, K.; Takiguchi, Y.
J. Biol. Chem. 263, 1231-1239, 1988
A:Title: Identification of a novel class of elastase isozyme, human pancreatic elastase
A:Reference number: A93664; PMID:88087253; PMID:2826474

Query Match 34.4%; Score 420.5; DB 2; Length 270;
Best Local Similarity 40.7%; Pred. No. 3.6e-28;
Matches 98; Conservative 35; Mismatches 91; Indels 17; Gaps 8;

[illegible]

Search completed: December 20, 2002, 15:19:47
Job time : 15.6087 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:14:53 : Search time 13.087 seconds
(without alignments)
519.349 Million cell updates/sec

Title: US-09-856-319B-4_COPY_34_264

Perfect score: 1252

Sequence: 1 IVNGENAVPGSWPQVSLQD.....AMTVRSKFTWVQVWVAYN 231

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	750.5	59.9	230	4	US-08-944-483-62
2	750.5	59.9	231	2	US-09-027-337-6
3	750.5	59.9	231	4	US-09-644-600-6
4	737	58.9	229	2	US-08-557-146-13
5	737	58.9	229	2	US-09-154-344-13
6	724.5	57.9	228	1	US-08-278-091-10
7	724.5	57.9	228	1	US-08-483-859-10
8	724.5	57.9	228	1	US-08-472-173-10
9	724.5	57.9	228	2	US-08-487-167-10
10	724.5	57.9	228	2	US-08-482-816-10
11	724.5	57.9	228	2	US-08-296-149-10
12	724.5	57.9	228	2	US-08-801-499-10
13	724.5	57.9	228	2	US-08-615-271-10
14	724.5	57.9	228	3	US-09-074-660-10
15	724.5	57.9	228	3	US-09-074-659-10
16	724.5	57.9	228	3	US-09-106-468-10
17	724.5	57.9	228	4	US-09-106-466A-10
18	724.5	57.9	228	4	US-09-106-467-10
19	486	38.8	241	4	US-08-944-483-59
20	481.5	38.5	242	4	US-08-944-483-57
21	466.5	37.3	814	1	US-08-750-711-1
22	465.5	37.2	791	1	US-08-643-219-1
23	465.5	37.2	791	3	US-08-851-350-1
24	464.5	37.1	230	1	US-08-456-840-47
25	464.5	37.1	230	1	US-08-266-407A-47
26	464.5	37.1	230	2	US-08-892-544-47
27	464.5	37.1	230	2	US-08-766-982-12

28	464.5	37.1	230	4	US-08-944-483-53
29	464.5	37.1	230	4	US-09-296-219-12
30	464.5	37.1	546	6	5200340-6
31	464.5	37.1	790	1	US-08-469-486-54
32	464.5	37.1	790	2	US-08-469-658-54
33	464.5	37.1	791	2	US-09-131-995-1
34	464.5	37.1	791	2	US-08-832-087B-1
35	464.5	37.1	791	4	US-09-132-154-1
36	464.5	37.1	810	1	US-07-854-603-2
37	464.5	37.1	810	1	US-08-147-000B-29
38	464.5	37.1	810	4	US-09-086-514-1
39	464.5	37.1	810	6	5200340-8
40	462	36.9	812	1	US-08-248-629A-1
41	462	36.9	812	1	US-08-451-932-1
42	462	36.9	812	1	US-08-452-260-1
43	462	36.9	812	1	US-08-326-785-1
44	462	36.9	812	2	US-08-612-788-1
45	462	36.9	812	2	US-08-605-598B-1

ALIGNMENTS

RESULT 1

US-08-944-483-62
: Sequence 62, Application US/08944483
: Patent No. 6232456

: GENERAL INFORMATION:

: APPLICANT: COHEN, MAURICE

: APPLICANT: COLPITTS, TRACEY L.

: APPLICANT: FRIEDMAN, PAULA N.

: APPLICANT: GRANADOS, EDWARD N.

: APPLICANT: KLASS, MICHAEL R.

: APPLICANT: RUSSELL, JOHN C.

: APPLICANT: STEWART, KENT D.

: APPLICANT: STROUPE, STEVEN D.

: TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS

: TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES

: TITLE OF INVENTION: OF THE PROSTATE

: NUMBER OF SEQUENCES: 76

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Abbott Laboratories

: STREET: 100 Abbott Park Road

: CITY: Abbott Park

: STATE: IL

: COUNTRY: USA

: ZIP: 60064-3500

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Diskette

: COMPUTER: IBM Compatible

: OPERATING SYSTEM: DOS

: SOFTWARE: FastSeq for Windows Version 2.0

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/944.483

: FILING DATE:

: CLASSIFICATION: 424

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER:

: FILING DATE:

: ATTORNEY/AGENT INFORMATION:

: NAME: Becker, Cheryl L.

: REGISTRATION NUMBER: 35,441

: REFERENCE/DOCKET NUMBER: 6183.US.01

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: 847/935-1729

: TELEFAX: 847/938-2623

: TELEX:

: INFORMATION FOR SEQ ID NO: 62:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 230 amino acids

: TYPE: amino acid

: STRANDEDNESS: single

: TOPOLOGY: linear

```
; MOLECULE TYPE: No. 62324566
US-08-944-483-62

Query Match          59.9%; Score 750.5; DB 4; Length 230;
Best Local Similarity 58.0%; Pred. No. 4.3e-64;
Matches 134; Conservative 35; Mismatches 61; Indels 1; Gaps 1;

QY 1 IVGENAVPGSWPQVSLQDNTGFHFCGGLISPNMVVTAACHQVTPGRHFVVLGEYDRS 60
Db 1 IVGEDAVPGSWPQVSLQDNTGFHFCGGLISPNMVVTAACHQVTPGRHFVVLGEYDRS 59
QY 61 SNAEPVOVLISARAIHPNNANTMNDLTLKLASPARYTAOVSPVCLASTNEALPSGL 120
Db 60 SDENIQVLKIAKVFKNPKFSILTNNITLLKLATPARFSQTVSAVCLPSADDDFPAGT 119
QY 121 TCVTGGRISGVNTVPARLQVPLVTVNOCROYWGARTIDAMICAGGSGASSCOGD 180
Db 120 LCATTGKTKYNANKTPDKLQQAALPLLSNAECKSKWGRRIIDVVICAGSGVSSCMGD 179
QY 181 SGGPLVCQKGNWTWLVIGISWGTKNQNIQAPAMYTRVSKFSTWVNOVMAYN 231
Db 180 SGGPLVCQKGNWTWLVIGISWGTKNQNIQAPAMYTRVSKFSTWVNOVMAYN 230

RESULT 2
US-09-027-337-6
; Sequence 6, Application US/09027337B
; Patent No. 5972616
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotochi
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease Overexpressed in
; FILE REFERENCE: D6064
; CURRENT FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 13
; SEQ ID NO 6
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Serine protease catalytic domain of chymotrypsin (Chymb)
; OTHER INFORMATION: homologous to similar domain in TADG-15
US-09-027-337-6

Query Match          59.9%; Score 750.5; DB 2; Length 231;
Best Local Similarity 58.0%; Pred. No. 4.3e-64;
Matches 134; Conservative 35; Mismatches 61; Indels 1; Gaps 1;

QY 1 IVGENAVPGSWPQVSLQDNTGFHFCGGLISPNMVVTAACHQVTPGRHFVVLGEYDRS 60
Db 2 IVGEDAVPGSWPQVSLQDNTGFHFCGGLISPNMVVTAACHQVTPGRHFVVLGEYDRS 60
QY 61 SNAEPVOVLISARAIHPNNANTMNDLTLKLASPARYTAOVSPVCLASTNEALPSGL 120
Db 61 SDENIQVLKIAKVFKNPKFSILTNNITLLKLATPARFSQTVSAVCLPSADDDFPAGT 120
QY 121 TCVTGGRISGVNTVPARLQVPLVTVNOCROYWGARTIDAMICAGGSGASSCOGD 180
Db 120 LCATTGKTKYNANKTPDKLQQAALPLLSNAECKSKWGRRIIDVVICAGSGVSSCMGD 180
QY 181 SGGPLVCQKGNWTWLVIGISWGTKNQNIQAPAMYTRVSKFSTWVNOVMAYN 231
Db 180 SGGPLVCQKGNWTWLVIGISWGTKNQNIQAPAMYTRVSKFSTWVNOVMAYN 231

RESULT 3
US-09-644-600-6
; Sequence 6, Application US/09644600
; Patent No. 6451500
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
```

```
; APPLICANT: Tanimoto, Hirotochi
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease
; FILE REFERENCE: D6064CIP/D
; CURRENT APPLICATION NUMBER: US/09/644,600
; CURRENT FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/421,213
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: 09/027,337
; PRIOR FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 6
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Chymotrypsin
US-09-644-600-6

Query Match          59.9%; Score 750.5; DB 4; Length 231;
Best Local Similarity 58.0%; Pred. No. 4.3e-64;
Matches 134; Conservative 35; Mismatches 61; Indels 1; Gaps 1;

QY 1 IVGENAVPGSWPQVSLQDNTGFHFCGGLISPNMVVTAACHQVTPGRHFVVLGEYDRS 60
Db 2 IVGEDAVPGSWPQVSLQDNTGFHFCGGLISPNMVVTAACHQVTPGRHFVVLGEYDRS 60
QY 61 SNAEPVOVLISARAIHPNNANTMNDLTLKLASPARYTAOVSPVCLASTNEALPSGL 120
Db 61 SDENIQVLKIAKVFKNPKFSILTNNITLLKLATPARFSQTVSAVCLPSADDDFPAGT 120
QY 121 TCVTGGRISGVNTVPARLQVPLVTVNOCROYWGARTIDAMICAGGSGASSCOGD 180
Db 120 LCATTGKTKYNANKTPDKLQQAALPLLSNAECKSKWGRRIIDVVICAGSGVSSCMGD 180
QY 181 SGGPLVCQKGNWTWLVIGISWGTKNQNIQAPAMYTRVSKFSTWVNOVMAYN 231
Db 180 SGGPLVCQKGNWTWLVIGISWGTKNQNIQAPAMYTRVSKFSTWVNOVMAYN 231

RESULT 4
US-08-557-146-13
; Sequence 13, Application US/08557146
; Patent No. 5834290
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hansson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,146
; FILING DATE: 14-DEC-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
```


; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 229 amino acids
; TYPE: amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
US-08-557-146-13

Query Match 58.9%; Score 737; DB 2; Length 229;
Best Local Similarity 57.6%; Pred. No. 8.3e-63;
Matches 133; Conservative 36; Mismatches 60; Indels 2; Gaps 2;
Qy 1 IVNGENAVPGSWPMQVSLQDNTGPHFCGSLISPNWVYTAACHCVTPGRHFVVLGEYDRS 60
Db 1 IVNGEDAVPGSWPMQVSLQDNTGPHFCGSLISEDWVYTAACHGVRTS-DVVVAGEFDQG 59
Qy 61 SNAEPVQVLSIARATHPHNWNANTMNDLTLKLLASPARYTAQVSPVCLASTNEALPSGL 120
Db 60 SDEENIQVLKIAKVFKNPKFSILTVNNDITLLKLTATPARFSQTVSAVCLPSADDDFPAGT 119
Qy 121 TCVTTGWRIGRISGVNVTAPRLOQVVLPLVTYNQCRQYWGARTIDAMICAGSGGSSCGD 180
Db 120 LCATTGWGKTKYNANKTPDKLQQAALPLLSNAECKKSGRRITDYMICA-GAGVSSCMGD 178
Qy 181 SGGPLVCOKGNTWVLIGIVSMGCTKNCNIQAPAMYTRVSKFSTWVNOVMAYN 231
Db 179 SGGPLVCOKGNTWVLIGIVSMGSDTCSTSSPGVYARVTKLIPWQKILAAAN 229

RESULT 5
US-09-154-344-13
; Sequence 13, Application US/09154344
; Patent No. 5981256
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hansson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/154,344
; FILING DATE: 16-SEP-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/557,146
; FILING DATE: 14-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-181
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 229 amino acids
; TYPE: amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: polypeptide
US-09-154-344-13

Query Match 58.9%; Score 737; DB 2; Length 229;
Best Local Similarity 57.6%; Pred. No. 8.3e-63;
Matches 133; Conservative 36; Mismatches 60; Indels 2; Gaps 2;
Qy 1 IVNGENAVPGSWPMQVSLQDNTGPHFCGSLISPNWVYTAACHCVTPGRHFVVLGEYDRS 60
Db 1 IVNGEDAVPGSWPMQVSLQDNTGPHFCGSLISEDWVYTAACHGVRTS-DVVVAGEFDQG 59
Qy 61 SNAEPVQVLSIARATHPHNWNANTMNDLTLKLLASPARYTAQVSPVCLASTNEALPSGL 120
Db 60 SDEENIQVLKIAKVFKNPKFSILTVNNDITLLKLTATPARFSQTVSAVCLPSADDDFPAGT 119
Qy 121 TCVTTGWRIGRISGVNVTAPRLOQVVLPLVTYNQCRQYWGARTIDAMICAGSGGSSCGD 180
Db 120 LCATTGWGKTKYNANKTPDKLQQAALPLLSNAECKKSGRRITDYMICA-GAGVSSCMGD 178
Qy 181 SGGPLVCOKGNTWVLIGIVSMGCTKNCNIQAPAMYTRVSKFSTWVNOVMAYN 231
Db 179 SGGPLVCOKGNTWVLIGIVSMGSDTCSTSSPGVYARVTKLIPWQKILAAAN 229

RESULT 6
US-08-278-091-10
; Sequence 10, Application US/08278091
; Patent No. 5506139
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; TITLE OF INVENTION: Reduced Protease Activity
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/278,091
; FILING DATE: 21-JUL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-371
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-278-091-10

Query Match 57.9%; Score 724.5; DB 1; Length 228;
Best Local Similarity 57.1%; Pred. No. 1.3e-61;
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;
Qy 1 IVNGENAVPGSWPMQVSLQDNTGPHFCGSLISPNWVYTAACHCVTPGRHFVVLGEYDRS 60

Db 1 IVNGEAVPGSWPQVSLQDKTGFFHFCGSLINENWVTAACHGVTTSDVYVAGEFDQ 59
QY 61 SNAEPQVLSIARATHPNWNTNNDTLTKLASPARYTAQVSPVCLASTNEALPSGL 120
Db 60 SSSERIKQIKIAKVFKNKSYNSLTINNDITLKLSTAAFSQTSVAVCLPSASDDFAAGT 119
QY 121 TCVTTGWGRISGVNTPARLQOVVPLVTVVNOCRQYWGARTDAMICAGGSGSCQGD 180
Db 120 TCVTTGWG-LTRYAN-TPDRLOQASLPPLSNNTNCKKYWGTKIKDAMICAGASGVSSCMGD 177
QY 181 SGGPLVCQKGNWTWLGIVSGTKNCNIOAPAMYTRVSKFSTWVINOVMAYN 231
Db 178 SGGPLVCKNGAWTLVGIVSGSSTCSTPGYARVATLVNMQVQTLAAN 228
RESULT 7
US-08-483-859-10
; Sequence 10, Application US/08483859
; Patent No. 5656436
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; NUMBER OF INVENTIONS: Reduced Protease Activity
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,859
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-495 MIS:vg
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-483-859-10
Query Match 57.9%; Score 724.5; DB 1; Length 228;
Best Local Similarity 57.1%; Pred. No. 1.3e-61;
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;

QY 1 IVNGEAVPGSWPQVSLQDKTGFFHFCGSLINENWVTAACHGVTTSDVYVAGEFDQ 59
Db 1 IVNGEAVPGSWPQVSLQDKTGFFHFCGSLINENWVTAACHGVTTSDVYVAGEFDQ 59

QY 61 SNAEPQVLSIARATHPNWNTNNDTLTKLASPARYTAQVSPVCLASTNEALPSGL 120
Db 60 SSSERIKQIKIAKVFKNKSYNSLTINNDITLKLSTAAFSQTSVAVCLPSASDDFAAGT 119
QY 121 TCVTTGWGRISGVNTPARLQOVVPLVTVVNOCRQYWGARTDAMICAGGSGSCQGD 180
Db 120 TCVTTGWG-LTRYAN-TPDRLOQASLPPLSNNTNCKKYWGTKIKDAMICAGASGVSSCMGD 177
QY 181 SGGPLVCQKGNWTWLGIVSGTKNCNIOAPAMYTRVSKFSTWVINOVMAYN 231
Db 178 SGGPLVCKNGAWTLVGIVSGSSTCSTPGYARVATLVNMQVQTLAAN 228
RESULT 8
US-08-472-173-10
; Sequence 10, Application US/08472173
; Patent No. 5665353
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; NUMBER OF INVENTIONS: Reduced Protease Activity
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,173
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-493 MIS:vg
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-472-173-10
Query Match 57.9%; Score 724.5; DB 1; Length 228;
Best Local Similarity 57.1%; Pred. No. 1.3e-61;
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;

QY 1 IVNGEAVPGSWPQVSLQDKTGFFHFCGSLINENWVTAACHGVTTSDVYVAGEFDQ 59
Db 1 IVNGEAVPGSWPQVSLQDKTGFFHFCGSLINENWVTAACHGVTTSDVYVAGEFDQ 59
QY 61 SNAEPQVLSIARATHPNWNTNNDTLTKLASPARYTAQVSPVCLASTNEALPSGL 120

Db 60 SSSEKIQLKIAKFKNSKYNLTINNDITLLKLTAAASFQSVSAVCLPSASDDFAAGT 119
Qy 121 TCVTTGWRISGVGNVTTPARLQOVVLPVTVNQCROYWGARITDAMICAGSGASSCGD 180
Db 120 TCVTTGWR-LTRYAN-TPDRLOQASLPLLSNTNCKKYGWGTIKIDAMICAGASGVSSCMGD 177
Qy 181 SGGPLVCOKGNTWVLIGVSGTKNCNQAPAMYTRVSKFSTWVNOVMAYN 231
Db 178 SGGPLVCKKNGAWTLVIGVSGSSCTSTPGVYARVTVLNVWQOTLAAN 228

RESULT 9
US-08-487-167-10
: Sequence 10, Application US/08487167
: Patent No. 5869302
: GENERAL INFORMATION:
: APPLICANT: LOOSMORE, Sheena M.
: APPLICANT: YANG, Yan-Ping
: APPLICANT: CHONG, Pele
: APPLICANT: OOMEN, Raymond P.
: APPLICANT: KLEIN, Michel H.
: TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
: TITLE OF INVENTION: Reduced Protease Activity
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sim & McBurney
: STREET: Suite 701, 330 University Avenue
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5G 1R7

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/487.167
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435

: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/296,149
: FILING DATE: 26-AUG-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/278,091
: FILING DATE: 21-JUL-1994

: ATTORNEY/AGENT INFORMATION:
: NAME: Stewart, Michael I.
: REGISTRATION NUMBER: 24,973
: REFERENCE/DOCKET NUMBER: 1038-508 MIS:vg
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (416) 595-1155
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 228 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear

US-08-487-167-10
Query Match 57.9%; Score 724.5; DB 2; Length 228;
Best Local Similarity 57.1%; Pred. No. 1.3e-61;
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;

Qy 1 IVNGENAVPGSPWQVSLQDNTGFHFCGSLSPNNWVTAACQVTPGRHFVVLGEYDRS 60
Db 1 IVNGEAVPGSPWQVSLQDNTGFHFCGSLSPNNWVTAACQVTPGRHFVVLGEYDRS 59
Qy 61 SNAEPQVLSIARATTHPNWNTNNDITLLKLTAAASFPQSVCLASTNEALPSGL 120
Db 60 SSSEKIQLKIAKFKNSKYNLTINNDITLLKLTAAASFPQSVCLASTNEALPSGL 119

Qy 121 TCVTTGWRISGVGNVTTPARLQOVVLPVTVNQCROYWGARITDAMICAGSGASSCGD 180
Db 120 TCVTTGWR-LTRYAN-TPDRLOQASLPLLSNTNCKKYGWGTIKIDAMICAGASGVSSCMGD 177
Qy 181 SGGPLVCOKGNTWVLIGVSGTKNCNQAPAMYTRVSKFSTWVNOVMAYN 231
Db 178 SGGPLVCKKNGAWTLVIGVSGSSCTSTPGVYARVTVLNVWQOTLAAN 228

RESULT 10
US-08-482-816-10
: Sequence 10, Application US/08482816
: Patent No. 5935573
: GENERAL INFORMATION:
: APPLICANT: LOOSMORE, Sheena M
: APPLICANT: YANG, Yan-Ping
: APPLICANT: CHONG, Pele
: APPLICANT: OOMEN, Raymond P.
: APPLICANT: KLEIN, Michel H.
: TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with Reduced Protease A
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sim & McBurney
: STREET: Suite 701, 330 University Avenue
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5G 1R7

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/482.816
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435

: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/296,149
: FILING DATE: 26-AUG-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/278,091
: FILING DATE: 21-JUL-1994

: ATTORNEY/AGENT INFORMATION:
: NAME: Stewart, Michael I.
: REGISTRATION NUMBER: 24,973
: REFERENCE/DOCKET NUMBER: 1038-494 MIS:vg
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (416) 595-1155
: TELEFAX: (416) 595-1163
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 228 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear

US-08-482-816-10
Query Match 57.9%; Score 724.5; DB 2; Length 228;
Best Local Similarity 57.1%; Pred. No. 1.3e-61;
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;

Qy 1 IVNGENAVPGSPWQVSLQDNTGFHFCGSLSPNNWVTAACQVTPGRHFVVLGEYDRS 60
Db 1 IVNGEAVPGSPWQVSLQDNTGFHFCGSLSPNNWVTAACQVTPGRHFVVLGEYDRS 59
Qy 61 SNAEPQVLSIARATTHPNWNTNNDITLLKLTAAASFPQSVCLASTNEALPSGL 120
Db 60 SSSEKIQLKIAKFKNSKYNLTINNDITLLKLTAAASFPQSVCLASTNEALPSGL 119
Qy 121 TCVTTGWRISGVGNVTTPARLQOVVLPVTVNQCROYWGARITDAMICAGSGASSCGD 180
Db 120 TCVTTGWR-LTRYAN-TPDRLOQASLPLLSNTNCKKYGWGTIKIDAMICAGASGVSSCMGD 177

Db 120 TCVTGNG-LTRYAN-TPDLQOASLPLSLNTNCKYWGTKIKDAMICAGASGVSSCMGD 177
QY 181 SGGPLVCQKGNWTWLVIGISWGTKNQCIQAPAMYTRVSKFSTWVNOVMAYN 231
Db 178 SGGPLVCCKNGAWTLVIGISWGSTCTSTPGVYARVATLVNWNVQOTLAAN 228

RESULT 11
US-08-296-149-10
; Sequence 10, Application US/08296149
; Patent No. 5939297
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,149
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-390
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-296-149-10

Query Match 57.9%; Score 724.5; DB 2; Length 228;
Best Local Similarity 57.1%; Pred. No. 1.3e-61;
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;

QY 1 IVGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAHQCQVTPGRHFVVLGEYDRS 60
Db 1 IVGEEAVPGSWPQVSLQDNTGFHFCGSLINENWVYTAHCGVTTs-DVVVAGEFDQ 59

QY 61 SNAEPQVLSIARAIHPNNANTMNDLTLKIASPARVTAQVSPVCLASTNEALPSGL 120
Db 60 SSSEKIQKIAKVFKNKNSYLSLTINNDITLLKLSAASFQTVSAVCLPSADDFAAGT 119

QY 121 TCVTGNGRISGVNTPPARLQOVLPVTVNOCROYWGARTIDAMICAGSGASSCOGD 180
Db 120 TCVTGNG-LTRYAN-TPDLQOASLPLSLNTNCKYWGTKIKDAMICAGASGVSSCMGD 177

QY 181 SGGPLVCQKGNWTWLVIGISWGTKNQCIQAPAMYTRVSKFSTWVNOVMAYN 231
Db 178 SGGPLVCCKNGAWTLVIGISWGSTCTSTPGVYARVATLVNWNVQOTLAAN 228

RESULT 12

US-08-801-499-10
; Sequence 10, Application US/08801499
; Patent No. 5962430
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with Reduced Protease A
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,499
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/482,816
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-671 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-801-499-10

Query Match 57.9%; Score 724.5; DB 2; Length 228;
Best Local Similarity 57.1%; Pred. No. 1.3e-61;
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;

QY 1 IVGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAHQCQVTPGRHFVVLGEYDRS 60
Db 1 IVGEEAVPGSWPQVSLQDNTGFHFCGSLINENWVYTAHCGVTTs-DVVVAGEFDQ 59

QY 61 SNAEPQVLSIARAIHPNNANTMNDLTLKIASPARVTAQVSPVCLASTNEALPSGL 120
Db 60 SSSEKIQKIAKVFKNKNSYLSLTINNDITLLKLSAASFQTVSAVCLPSADDFAAGT 119

QY 121 TCVTGNGRISGVNTPPARLQOVLPVTVNOCROYWGARTIDAMICAGSGASSCOGD 180
Db 120 TCVTGNG-LTRYAN-TPDLQOASLPLSLNTNCKYWGTKIKDAMICAGASGVSSCMGD 177

QY 181 SGGPLVCQKGNWTWLVIGISWGTKNQCIQAPAMYTRVSKFSTWVNOVMAYN 231
Db 178 SGGPLVCCKNGAWTLVIGISWGSTCTSTPGVYARVATLVNWNVQOTLAAN 228

RESULT 13

US-08-615-271-10
; Sequence 10, Application US/08615271
; Patent No. 5981503
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/615,271
; FILING DATE: 20-JUN-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-580
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-615-271-10

Query Match 57.9%; Score 724.5; DB 2; Length 228;
Best Local Similarity 57.1%; Pred. No. 1.3e-61;
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;
Qy 1 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAHQCQVTPGRHFVVLGEYDRS 60
Db 1 IVNGEEAVPGSWPQVSLQDNTGFHFCGSLINENWVYTAHCGVTS-DVYVAGEFDQG 59
Qy 61 SNAEPVQVLSARATHPHNWNANTMNDLTLKLSAPARYTAQVSPVCLASTNEALPSGL 120
Db 60 SSSEKIQKLKIAKVPKNSKYNLSLTINDITLLKLSAASFQSVAVCLPSASDDFAAGT 119
Qy 121 TCVTTGWRISGVGNVTPARLQVVLPLVTNQCQRYNGARITDAMICAGSGSSCCGD 180
Db 120 TCVTTGWL-LTRYAN-TPDRLQOASLPLLSNTNCKYNGTKIKDAMICAGSGVSSCMGD 177
Qy 181 SGGPLVCQKGNWTVLIGVSGTKNCNTQAPAMYTRVSKFSTWVQVMAYN 231
Db 178 SGGPLVCKKNGAWTLVIGVSGSSCTSTPGVYARVYALVNWVQOTLAAN 228
RESULT 14
US-09-074-660-10
; Sequence 10, Application US/09074660
; Patent No. 6020183
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.

; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/074,660
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,167
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-731 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-074-660-10
Query Match 57.9%; Score 724.5; DB 3; Length 228;
Best Local Similarity 57.1%; Pred. No. 1.3e-61;
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;
Qy 1 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAHQCQVTPGRHFVVLGEYDRS 60
Db 1 IVNGEEAVPGSWPQVSLQDNTGFHFCGSLINENWVYTAHCGVTS-DVYVAGEFDQG 59
Qy 61 SNAEPVQVLSARATHPHNWNANTMNDLTLKLSAPARYTAQVSPVCLASTNEALPSGL 120
Db 60 SSSEKIQKLKIAKVPKNSKYNLSLTINDITLLKLSAASFQSVAVCLPSASDDFAAGT 119
Qy 121 TCVTTGWRISGVGNVTPARLQVVLPLVTNQCQRYNGARITDAMICAGSGSSCCGD 180
Db 120 TCVTTGWL-LTRYAN-TPDRLQOASLPLLSNTNCKYNGTKIKDAMICAGSGVSSCMGD 177
Qy 181 SGGPLVCQKGNWTVLIGVSGTKNCNTQAPAMYTRVSKFSTWVQVMAYN 231
Db 178 SGGPLVCKKNGAWTLVIGVSGSSCTSTPGVYARVYALVNWVQOTLAAN 228
RESULT 15
US-09-074-659-10
; Sequence 10, Application US/09074659
; Patent No. 6025342
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele

APPLICANT: OOMEN, Raymond P.
APPLICANT: KLEIN, Michel H.
TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
TITLE OF INVENTION: Reduced Protease Activity
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/074,659
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/487,167
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/296,149
FILING DATE: 26-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/278,091
FILING DATE: 21-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-730 MIS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 228 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-074-659-10

Query Match 57.9%; Score 724.5; DB 3; Length 228;
Best Local Similarity 57.1%; Pred. No. 1.3e-61;
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;
QY 1 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVTAACHQVTPGRHFVVLGEYDRS 60
Db 1 IVNGEAVPGSWPQVSLQDNTGFHFCGSLISPNWVTAACHQVTPGRHFVVLGEYDRS 60
QY 61 SNAEPQVLSIARATHPNMNTMNDITLLKLASPARYTAQVSPVCLASTNEALPSGL 120
Db 60 SSSEKIQLKIAKVFKNKYNLSITNNDITLLKLSTAAFSQTSVAVCLPSASDDFAAGT 119
QY 121 TCVTTGWRISGVNTPARLQOVLPVTVNOCROYGWRITDAMICAGSGSCCGD 180
Db 120 TCVTTGWRISGVNTPARLQOVLPVTVNOCROYGWRITDAMICAGSGSCCGD 177
QY 181 SGGPLVCKNGATLVIGIVSGWSSTCSTSTPGVYARVATLVNQQTLAAN 231
Db 178 SGGPLVCKNGATLVIGIVSGWSSTCSTSTPGVYARVATLVNQQTLAAN 228

Search completed: December 20, 2002, 15:20:39
Job time : 14.087 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:12:58 : Search time 7.91304 Seconds
(without alignments)
1210.790 Million cell updates/sec

Title: US-09-856-319B-4_COPY_34_264

Perfect score: 1252

Sequence: 1 IVNGENAVPGSWPQVSLQD.....AMYTRVSKFSTWVQVNAVYN 231

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1119	89.4	264	1	CTRL_HUMAN
2	757.5	60.5	263	1	CTRL_CANFA
3	750.5	59.9	263	1	CTRL_GADMO
4	750.5	59.9	263	1	CTRL_HUMAN
5	740.5	59.1	245	1	CTRL_BOVIN
6	727.5	58.1	245	1	CTRL_BOVIN
7	715.5	57.1	263	1	CTRL_RAT
8	714.5	57.1	245	1	CTRL_GADMO
9	503	40.2	269	1	EL2_PIG
10	499	39.9	271	1	EL2_RAT
11	497.5	39.7	253	1	CAC3_BOVIN
12	486	38.8	269	1	EL2A_HUMAN
13	481.5	38.5	270	1	EL3B_HUMAN
14	480	38.3	269	1	EL2_BOVIN
15	475	37.9	271	1	EL2_MOUSE
16	464.5	37.1	810	1	PLMN_HUMAN
17	462	36.9	812	1	PLMN_MOUSE
18	459.5	36.7	270	1	EL3A_HUMAN
19	455.5	36.4	270	1	TRYP_MERUN
20	455.5	36.4	342	1	PS8_RAT
21	455	36.3	454	1	TMS3_HUMAN
22	452.5	36.1	343	1	TRYP_PIG
23	452.5	36.1	343	1	PLMN_SHEEP
24	451.5	36.1	810	1	PLMN_MACMU
25	449.5	35.9	790	1	PLMN_PIG
26	449	35.9	269	1	EL2B_HUMAN
27	448.5	35.8	333	1	PLMN_CANFA
28	448.5	35.8	343	1	PS8_HUMAN
29	447.5	35.7	268	1	CLCR_HUMAN
30	447.5	35.7	268	1	CLCR_RAT
31	447.5	35.7	273	1	MCT7_MOUSE
32	447.5	35.7	342	1	PS8_MOUSE
33	444	35.5	231	1	TRYP_PIG

34	442.5	35.3	273	1	MCT7_RAT	P27435	rattus norv
35	440	35.1	246	1	TRV1_RAT	P00762	rattus norv
36	439.5	35.1	274	1	MCT6_RAT	P50343	rattus norv
37	439.5	35.1	490	1	TMS2_MOUSE	Q9J1Q8	mus musculu
38	439	35.1	248	1	TRV1_CHICK	Q90627	gallus gall
39	437.5	34.9	338	1	PLMN_HORSE	P80010	equus cabal
40	436	34.8	246	1	TRV1_CANFA	P06871	canis famil
41	435	34.7	248	1	TRV2_CHICK	Q90628	gallus gall
42	433	34.6	247	1	TRV2_CANFA	P06872	canis famil
43	431	34.4	246	1	TRV2_RAT	P00763	rattus norv
44	430.5	34.4	266	1	EL1_RAT	P00773	rattus norv
45	430.5	34.4	276	1	MCT6_MOUSE	P21845	mus musculu

ALIGNMENTS

RESULT 1				
ID	CTRL_HUMAN	STANDARD;	PRT;	264 AA.
AC	P40313;			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Chymotrypsin-like protease CTRL-1 precursor (EC 3.4.21.-).			
GN	CTRL OR CTRL1			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94093544; PubMed=8268911;			
RA	Larsen F., Solheim J., Kristensen T., Kolsto A.B., Prydz H.;			
RT	"A tight cluster of five unrelated human genes on chromosome 16q22.1."			
RL	Hum. Mol. Genet. 2:1589-1595(1993).			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			

CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	EMBL; X71874; CAA50710.1; -
DR	EMBL; X71877; CAA50711.1; -
DR	HSSP; P00763; IDPO.
DR	MEROPS; S01.256; -
DR	Genew; HGNC:2524; CTRL.
DR	MIM; 118888; -
DR	InterPro; IPR001314; Chymotrypsin.
DR	InterPro; IPR001254; Ser_protease_Try.
DR	Pfam; PF00089; trypsin; 1.
DR	PRINTS; PR00722; CHYMOTRYPSIN.
DR	SMART; SM00020; TRYP_SPC; 1.
DR	PROSITE; PS50240; TRYP_SIN_DOM; 1.
DR	PROSITE; PS00134; TRYP_SIN_HIS; 1.
DR	PROSITE; PS00135; TRYP_SIN_SER; 1.
KW	Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal.
FT	SIGNAL 1 18 POTENTIAL.
FT	PROPEP 19 33* ACTIVATION PEPTIDE (POTENTIAL).
FT	CHAIN 34 264 CHYMOTRYPSIN-LIKE PROTEASE CTRL-1.
FT	ACT_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 121 121 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 214 214 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID 19 141 BY SIMILARITY.
FT	DISULFID 60 76 BY SIMILARITY.
FT	DISULFID 155 220 BY SIMILARITY.
FT	DISULFID 187 201 BY SIMILARITY.

FT	CHAIN	34	164	CHYMOTRYPSIN 2, B CHAIN.
FT	CHAIN	167	263	CHYMOTRYPSIN 2, C CHAIN.
FT	ACT_SITE	75	75	CHARGE RELAY SYSTEM.
FT	ACT_SITE	120	120	CHARGE RELAY SYSTEM.
FT	ACT_SITE	213	213	CHARGE RELAY SYSTEM.
FT	DISULFID	19	140	BY SIMILARITY.
FT	DISULFID	60	76	BY SIMILARITY.
FT	DISULFID	154	219	BY SIMILARITY.
FT	DISULFID	186	200	BY SIMILARITY.
FT	DISULFID	209	238	BY SIMILARITY.
SEQ	SEQUENCE	263 AA;	27787 MW;	2A2F449D813B3961 CRC64;

Query Match 60.5%; Score 757.5; DB 1; Length 263;
Best Local Similarity 58.0%; Pred. No. 7.7e-61;
Matches 134; Conservative 35; Mismatches 61; Indels 1; Gaps

QY	1	IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVVTAAHCQVTPGRHFVVLGEBYDRS	60
DB	34	IVNGENAVPGSWPQVSLQDNTGFHFCGSLISEDWVVTAAHCQVTR-THQVVAGEFDG	92
QY	61	SNAPVQVLSIARAIHPNNNTMMNDLTLLKLSPARYTAQVSPVCLASTNEALPSGL	120
DB	93	SDAESTQVLKIAKFKPKFMFTIINDITLLKLTAPRFKSTVSAVCLPQATDDPACT	152
QY	121	TCVTTGWRISGVGNVTPARLOQVWLPLVTVNCQRYGARITDAMICAGGSGASSCGD	180
DB	153	LCVTTGWLTKHTNANTPDKLQAALPILLSNAECKKFWGSKITDLVMVCAGSGVSSCMGD	212
QY	181	SGGPLVCQKQNTWVLGIYVSWGTKNCNIOAPAMYTRVSKFSTWINOVMAYN	231
DB	213	SGGPLVCQKQDAGTLVGIYVSGTSGCTSTPGVYARVTKLIPWVQILOAN	263

RESULT 3

ID	CTRA_GADMO	STANDARD;	PRT;	263 AA.
AC	P47796;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Chymotrypsin A precursor (EC 3.4.21.1).			
OS	Gadus morhua (Atlantic cod).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;			
OC	Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.			
OX	NCBI_Taxid=8049;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Pyloric caeca;			
RX	MEDLINE=94368860; PubMed=8086467;			
RA	Gudmundsdottir A., Oskarsson S., Eakin A.E., Craik C.S.,			
RA	Bjarnason J.B.;			
RL	"Atlantic cod cDNA encoding a psychrophilic chymotrypsinogen.";			
RT	Biochim. Biophys. Acta 1219:211-214(1994).			
RN	[2]			
RP	SEQUENCE OF 19-30 AND 34-49.			
RC	TISSUE=Pyloric caeca;			
RX	MEDLINE=92111252; PubMed=1764912;			
RA	Asgerlsson B., Bjarnason J.B.;			
RT	"Structural and kinetic properties of chymotrypsin from Atlantic cod			
RT	(Gadus morhua). Comparison with bovine chymotrypsin."			
RL	Comp. Biochem. Physiol. 99B:327-335(1991).			
CC	-I- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,			
CC	Phe-I-Xaa, Leu-I-Xaa.			
CC	-I- SUBCELLULAR LOCATION: Extracellular.			
CC	-I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on the use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcements>).

or send an email to license@isb-sib.ch).

```

CC EMBL; X78490; CAA5242.1; -.
CC HSSP; P00766; ICHG.
CC MEROPS; S01.152; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
FT SIGNAL 1 18
FT CHAIN 19 263 CHYMOTRYPSIN A.
FT ACT_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 19 140 BY SIMILARITY.
FT DISULFID 60 76 BY SIMILARITY.
FT DISULFID 154 219 BY SIMILARITY.
FT DISULFID 186 200 BY SIMILARITY.
FT DISULFID 209 238 BY SIMILARITY.
FT CONFLICT 21 21 R -> S (IN REF. 2).
FT CONFLICT 25 25 S -> Q (IN REF. 2).
FT CONFLICT 29 29 T -> S (IN REF. 2).
FT CONFLICT 44 44 S -> T (IN REF. 2).
FT CONFLICT 46 46 S -> Y (IN REF. 2).
SQ SEQUENCE 263 AA; 28294 MW; 47AAC699A0A64FBB CRC64;

```

```

Query Match 59.9%; Score 750.5; DB 1; Length 263;
Best Local Similarity 58.4%; Pred. No. 3.3e-60;
Matches 135; Conservative 33; Mismatches 62; Indels 1; Gaps 1;

QY 1 IVNGENAVPGSWPQVSLQDNTGFHCGGLISPNWVYTAACHQVTPGRHFVVLGEYDYS 60
DB 34 IVNGEEAVPHSWQVSLQDQTFHFCGGLINENWVYTAACHNV-KNYHRVVLGEHDS 92
QY 61 SNAEPQVLSSTARAIHPNNANTMNDLTLKLLASPARYTAQVSPVCLASTNEALPSGL 120
DB 93 SNSEGVQVMTVGQVFKRHPNGFTINDILLVKLATPATLNMRVSPVCLAEITDDVFE 152
QY 121 TCVTGTGGRISGVGNVTPARLQVVLPLVTNOCROYGARGITDAMICAGSGASSCOGD 180
DB 153 KVTSGWGLTRYNAADTPALLQQAALPLLTNEQCKKFWGNKISDLMLCAGAGASSCMGD 212
QY 181 SGGPLVCQKGNWTVLIGIVSGWTKNCKNIQAPAMTRYSKFSTWYNQVAYN 231
DB 213 SGGPLVCQKAGSWTLVIGVSGWGTCTPTMPGVYARVTELRAWVDQTAAN 263

```

```

RESULT 4
CTRB_HUMAN STANDARD; PRT; 263 AA.
ID CTRB_HUMAN
AC P17538;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsinogen B precursor (EC 3.4.21.1).
GN CTRB1 OR CTRB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=89134264; PubMed=2917002;
RA Tomita N., Izumoto Y., Horii A., Doi S., Yokouchi H., Ogawa M.,
RA Mori T., Matsubara K.;
RT "Molecular cloning and nucleotide sequence of human pancreatic
prechymotrypsinogen cDNA.";

```

```

RL Biochem. Biophys. Res. Commun. 158:569-575(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
CC Phe-I-Xaa, Leu-I-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; M24400; AAA52128.1; -.
EMBL; BC005385; AAH05385.1; -.
DR PIR; A31299; A31299.
DR HSSP; P00766; ICHG.
DR MEROPS; S01.152; -.
DR Genew; HGNC:2521; CTRB1.
DR MIM; 118890; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
FT SIGNAL 1 18
FT CHAIN 19 263 CHYMOTRYPSIN B.
FT CHAIN 19 31 CHYMOTRYPSIN B, A CHAIN.
FT CHAIN 34 164 CHYMOTRYPSIN B, B CHAIN.
FT CHAIN 167 263 CHYMOTRYPSIN B, C CHAIN.
FT ACT_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 19 140 BY SIMILARITY.
FT DISULFID 60 76 BY SIMILARITY.
FT DISULFID 154 219 BY SIMILARITY.
FT DISULFID 186 200 BY SIMILARITY.
FT DISULFID 209 238 BY SIMILARITY.
SQ SEQUENCE 263 AA; 27870 MW; 4C1C055A490B8701 CRC64;

Query Match 59.9%; Score 750.5; DB 1; Length 263;
Best Local Similarity 58.0%; Pred. No. 3.3e-60;
Matches 134; Conservative 35; Mismatches 61; Indels 1; Gaps 1;

QY 1 IVNGENAVPGSWPQVSLQDNTGFHCGGLISPNWVYTAACHQVTPGRHFVVLGEYDYS 60
DB 34 IVNGEDAVPGSWPQVSLQDQTFHFCGGLISEDWVYTAACHGVRTS-DVVAGEFDQG 92
QY 61 SNAEPQVLSSTARAIHPNNANTMNDLTLKLLASPARYTAQVSPVCLASTNEALPSGL 120
DB 93 SDEENIQVLKIAKVFKNPKPFSILTAVNNDITLKLATPARFSQTVSAVCLPSADDDFPAGT 152
QY 121 TCVTGTGGRISGVGNVTPARLQVVLPLVTNOCROYGARGITDAMICAGSGASSCOGD 180
DB 153 LCATGTGKTKYNANKTPDKLQQAALPLLSNAECKSKGRRITDVMICAGASGVSSCMGD 212
QY 181 SGGPLVCQKGNWTVLIGIVSGWTKNCKNIQAPAMTRYSKFSTWYNQVAYN 231
DB 213 SGGPLVCQKDGAWTLVIGVSGWSDTCTSSSPGVYARVTKLIPVQKILAAAN 263

RESULT 5
CTRB_BOVIN

```

```
ID CTRB_BOVIN STANDARD; PRT; 245 AA.
AC P00767;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsinogen B (EC 3.4.21.1).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE, DISULFIDE BONDS, AND ACTIVE SITE.
RX MEDLINE=68238908; PubMed=5649671;
RA Smillie L.B., Furka A., Nagabhushan N., Stevenson K.J., Parkes C.O.;
RT "Structure of chymotrypsinogen B compared with chymotrypsinogen A and
trypsinogen.";
RL Nature 218:343-346(1968).
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
Phe-I-Xaa, Leu-I-Xaa.
CC -|- SUBCELLULAR LOCATION: Extracellular.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -|- DATABASE: NAME-Worthington enzyme manual;
CC WWW="http://www.worthington-biochem.com/manual/C/CHY.html".
DR PIR; A00953; KYBOB.
DR HSSP; P00766; IACB.
DR MEROPS; S01.152; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen.
FT CHAIN 1 13
FT CHAIN 16 146 CHYMOTRYPSIN B, A CHAIN.
FT CHAIN 149 245 CHYMOTRYPSIN B, B CHAIN.
FT ACT_SITE 57 57 CHYMOTRYPSIN B, C CHAIN.
FT ACT_SITE 102 102 CHARGE RELAY SYSTEM.
FT ACT_SITE 195 195 CHARGE RELAY SYSTEM.
FT DISULFID 1 122 CHARGE RELAY SYSTEM.
FT DISULFID 42 58
FT DISULFID 136 201
FT DISULFID 168 182
FT DISULFID 191 220
SQ SEQUENCE 245 AA; 25755 MW; 678016445FF5FEB5 CRC64;

Query Match 59.1%; Score 740.5; DB 1; Length 245;
Best Local Similarity 57.1%; Pred. No. 2.4e-59;
Matches 132; Conservative 32; Mismatches 66; Indels 1; Gaps 1;

QY 1 IYGENAVPGSPWQVSLQDNTGFHFCGSLISPNWVYTAACQVTPGRHFVVLGEYDRS 60
DB 16 IYNGDAVPGSPWQVSLQDSTGFHFCGSLISDWDVYTAACGVTTS-DVVVAGEFDQG 74
QY 61 SNAEPQVLSIARATHPNWNNANTMNDLTLKLASPARYTAQVSPVCLASTNEALPSGL 120
DB 75 LETEDQVQLKIGKVEKPKFSILTVRNDITLLKLATPAQSETVSACLPSADEDFPAGM 134
QY 121 TCVTYTGWRISGVNVTPARLOOVVLPLVTYVNCQRYQWGARITDAMICAGSGASCQGD 180
DB 135 LCATTGWGKTKYNALKTPDKLOQATLPIVSNPTDCRYKSGSRVTDVNICAGASGVSCMGD 194
QY 181 SGGPLVCOKGNTWVLGIVSWCKTNCNIOAPAMYTRVSKFSFWINQVMAYN 231
DB 195 SGGPLVCQKNGAWTLGIVSWGSTCSTSTPVAIYARVATLMPWQVETLAAN 245

RESULT 6
ID CTRB_BOVIN STANDARD; PRT; 245 AA.
```

```
AC P00766;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsinogen A (EC 3.4.21.1).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE, DISULFIDE BONDS, AND ACTIVE SITE.
RX MEDLINE=67181721; PubMed=5971783;
RA Brown J.R., Hartley B.S.;
RT "Location of disulphide bridges by diagonal paper electrophoresis.
The disulphide bridges of bovine chymotrypsinogen A.";
RL Biochem. J. 101:214-228(1966).
RN [2]
RP REVISION TO 102.
RX MEDLINE=69106266; PubMed=5764436;
RA Blow D.M., Birktoft J.J., Hartley B.S.;
RT "Role of a buried acid group in the mechanism of action of
chymotrypsin.";
RL Nature 221:337-340(1969).
RN [3]
RP PRELIMINARY SEQUENCE.
RA Hartley B.S.;
RT "Amino-acid sequence of bovine chymotrypsinogen-A.";
RL Nature 201:1284-1287(1964).
RN [4]
RP SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=67183948; PubMed=5972866;
RA Meloun B., Kluh I., Kostka V., Moravek L., Prusik Z., Vanacek J.,
Keil B., Sorm F.;
RT "Covalent structure of bovine chymotrypsinogen A.";
RL Biochim. Biophys. Acta 130:543-546(1966).
RN [5]
RP ACTIVE SITE.
RX MEDLINE=67181723; PubMed=5971785;
RA Smillie L.B., Hartley B.S.;
RT "Histidine sequences in the active centres of some 'serine'
proteinases.";
RL Biochem. J. 101:232-241(1966).
RN [6]
RP X-RAY CRYSTALLOGRAPHY.
RX MEDLINE=72035052; PubMed=4399050;
RA Birktoft J.J., Blow D.M., Henderson R., Steitz T.A.;
RT "I. Serine proteinases. The structure of alpha-chymotrypsin.";
RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 257:76-76(1970).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF CHYMOTRYPSIN.
RX MEDLINE=70177557; PubMed=5442169;
RA Freer S.T., Kraut J., Robertus J.D., Wright H.T., Xuong N.H.;
RT "Chymotrypsinogen: 2.5-A crystal structure, comparison with alpha-
chymotrypsin, and implications for zymogen activation.";
RL Biochemistry 9:1997-2009(1970).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF GAMMA-CHYMOTRYPSIN.
RX MEDLINE=82078042; PubMed=6914398;
RA Cohen G.H., Silverton E.W., Davies D.R.;
RT "Refined crystal structure of gamma-chymotrypsin at 1.9-A resolution.
Comparison with other pancreatic serine proteases.";
RL J. Mol. Biol. 148:449-479(1981).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (1.68 ANGSTROMS) OF ALPHA-CHYMOTRYPSIN.
RX MEDLINE=86011575; PubMed=4046030;
RA Tsukada H., Blow D.M.;
RT "Structure of alpha-chymotrypsin refined at 1.68-A resolution.";
RL J. Mol. Biol. 184:703-711(1985).
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
Phe-I-Xaa, Leu-I-Xaa.
CC -|- SUBCELLULAR LOCATION: Extracellular.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
```



```
RL J. Biol. Chem. 259:14265-14270(1984)).
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
CC Phe-I-Xaa, Leu-I-Xaa.
CC -|- SUBCELLULAR LOCATION: Extracellular.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; K02298; AAA98732.1; -.
CC PIR; A22658; KYRTB.
CC HSP; P00766; ICHG.
CC MEROPS; S01.152; -.
CC InterPro; IPR001314; Chymotrypsin.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC SMART; SM00020; TRYPSIN_DOM; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
FT SIGNAL 1 18
FT CHAIN 19 263 CHYMOTRYPSINOGEN B.
FT CHAIN 19 31 CHYMOTRYPSIN B, A CHAIN.
FT CHAIN 34 164 CHYMOTRYPSIN B, B CHAIN.
FT CHAIN 167 263 CHYMOTRYPSIN B, C CHAIN.
FT ACT_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 19 140 BY SIMILARITY.
FT DISULFID 60 76 BY SIMILARITY.
FT DISULFID 154 219 BY SIMILARITY.
FT DISULFID 186 200 BY SIMILARITY.
FT DISULFID 209 238 BY SIMILARITY.
SQ SEQUENCE 263 AA; 27849 MW; ACAFD8ACF8C4DA6D CRC64;

Query Match 57.1%; Score 715.5; DB 1; Length 263;
Best Local Similarity 55.4%; Pred. No. 4.5e-57;
Matches 128; Conservative 35; Mismatches 67; Indels 1; Gaps 1;

QY 1 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVTAACHCQVTPGRHFFVLGEYDRS 60
Db 34 IVNGEDAIPGSPWQVSLQDNTGFHFCGSLISEDWVTAACHGVKTS-DVVVAGEFDQG 92

QY 61 SNAEPQVLSIARATTHPNWNTANNNDTLTKLASPARYTAQVSPVCLASTNEALPSGL 120
Db 93 SDEENIQVLKIAQVFKPNFNNFTVRNDITLLKATPAQFSETVSACVCLPNVDDDFPPGT.152

QY 121 TCVTTGWGRISGVNVTAPRLOQVPLPLVTNQCQRYWGARITDAMICAGGSGASCCGD 180
Db 153 VCATTGWGKTKYNAUKTEPKLQQAALPIVSEADCKKSGSKITDYMTCAGASGVSCMGD 212

QY 181 SGGPLVCQKGNVTWLVIGVSWGKTKNCIQAPAMYTRVSKFSTWVNOVMAYN 231
Db 213 SGGPLVCQKDGWVTLGIVSWGSGVCSITSPAYISRVTALMPWVOQILEAN 263

RESULT 8
CTRB_GADMO STANDARD; PRT; 245 AA.
AC P80646;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsin B (EC 3.4.21.1).
OS Gadus morhua (Atlantic cod).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
OX NCBI_TaxID=8049;
RN [1]
RC SEQUENCE.
RC TISSUE=Pyloric caeca;
RX MEDLINE=96439045; PubMed=8841380;
RA Leth-Larsen R., Asgeirsson B., Thorolfsson M., Noerregaard-Madsen M.,
RA Hoejrup P.;
RT "Structure of chymotrypsin variant B from Atlantic cod, Gadus
RT morhua.";
RL Biochim. Biophys. Acta 1297:49-56(1996).
RN [2]
RP SEQUENCE OF 1-12 AND 16-31.
RC TISSUE=Pyloric caeca;
RX MEDLINE=92111252; PubMed=1764912;
RA Asgeirsson B., Bjarnason J.B.;
RT "Structural and kinetic properties of chymotrypsin from Atlantic cod
RT (Gadus morhua). Comparison with bovine chymotrypsin.";
RL Comp. Biochem. Physiol. 99B:327-335(1991).
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
CC Phe-I-Xaa, Leu-I-Xaa.
CC -|- SUBCELLULAR LOCATION: Extracellular.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC HSP; P00766; ICHG.
CC MEROPS; S01.152; -.
CC InterPro; IPR001314; Chymotrypsin.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC SMART; SM00020; TRYPSIN_DOM; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen.
FT CHAIN 1 13 CHYMOTRYPSIN B, A CHAIN.
FT CHAIN 16 245 CHYMOTRYPSIN B, B CHAIN.
FT ACT_SITE 57 57 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 101 101 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 195 195 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 1 121 BY SIMILARITY.
FT DISULFID 42 58 BY SIMILARITY.
FT DISULFID 135 201 BY SIMILARITY.
FT DISULFID 167 182 BY SIMILARITY.
FT DISULFID 191 220 BY SIMILARITY.
FT CONFLICT 9 11 QVT -> VIS (IN REF. 2).
FT CONFLICT 26 26 S -> T (IN REF. 2).
FT CONFLICT 28 29 PW -> Y (IN REF. 2).
SQ SEQUENCE 245 AA; 26260 MW; 74FE0D425517AB02 CRC64;

Query Match 57.1%; Score 714.5; DB 1; Length 245;
Best Local Similarity 55.1%; Pred. No. 5.1e-57;
Matches 129; Conservative 41; Mismatches 57; Indels 7; Gaps 5;

QY 1 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVTAACHCQVTPGRHFFVLGEYDRS 60
Db 16 IVNGEAVPHSPWQVSLQDNTGFHFCGSLISPNWVTAACHCQVTPGRHFFVLGEYDRS 74

QY 61 SNAEPQVLSIARATTHPNWNTANNNDTLTKLASPARYTAQVSPVCLASTNEALPSGL 120
Db 75 SD-ENIQILKPSWVTHPKWDSRTINNDLSLTKLASPAVLGTNVPVCLGESSDVFAPGM 133

QY 121 TCVTTGWG--RISGVNVTAPRLOQVPLPLVTNQCQRYWGAR--ITDAMICAGGSGASCC 177
Db 134 KCVTSWGLTRYNAPG--TPNKLQQAALPLMSNECSQTNWNNMISDVWICAGAGATSC 191

QY 178 QDGGGLVLCQKGNVTWLVIGVSWGKTKNCIQAPAMYTRVSKFSTWVNOVMAYN 231
Db 192 MGDGGGLVLCQKGNVTWLVIGVSWGSSRCSTVTPAVYARVTELRLGWVQDQILAN 245

RESULT 9
EL2_PIG
```

```

ID EL2_PIG STANDARD; PRT; 269 AA.
AC P08419;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Elastase 2 precursor (EC 3.4.21.71).
GS ELA2.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87217962; PubMed=3646943;
RA Kawashima I., Tani T., Shimoda K., Takiguchi Y.;
RA Shimada Y., Ozawa T., Mikayama T., Iijima H., Ishida A., Sato Y.,
RA Tanai Y., Tanaka J., Ikenaga H.;
RT "Molecular cloning and expression in Escherichia coli of a cDNA
RT encoding human pancreatic elastase 2.";
RL J. Biochem. 102:1555-1563(1987).
CC -|- FUNCTION: ACTS UPON ELASTIN.
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Leu-|-Xaa, Met-|-Xaa
CC and Phe-|-Xaa. Hydrolyzes elastin.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- TISSUE SPECIFICITY: PANCREAS.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M16651; AAA31027.1; -.
DR EMBL; D00237; BAA00166.1; -.
DR PIR; A26823; A26823.
DR HSP; P00763; LDPO.
DR MEROPS; S01.155; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Src; 1.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolyase; Serine protease; zymogen; Signal.
FT SIGNAL 1 16
FT PROPEP 17 28
FT CHAIN 29 269
FT ACT_SITE 73 73
FT ACT_SITE 121 121
FT ACT_SITE 216 216
FT ACT_SITE 216 216
FT DISULFID 58 74
FT DISULFID 155 222
FT DISULFID 186 202
FT DISULFID 212 243
FT CONFLICT 10 10
FT CONFLICT 118 118
FT CONFLICT 132 132
FT CONFLICT 172 172
FT CONFLICT 202 202
FT CONFLICT 269 AA; 28699 MW; BAC6FE59AF4DDE56 CRC64;
SQ SEQUENCE 269 AA; 28699 MW; BAC6FE59AF4DDE56 CRC64;

Query Match 40.2%; Score 503; DB 1; Length 269;
Best Local Similarity 44.2%; Pred. No. 5, 1e-38;
Matches 107; Conservative 41; Mismatches 82; Indels 12; Gaps 9;

QY 1 IVNGENAVPGSWPQVOVSLQ-DNTG--FHPCGSLSPNNVVTAACHQVTPGRRHFVVLGEY 57
DB :|||:| ||||| |||:| |||:| |||:| |||:| |||:| |||:| |||:|
DB 29 VYGGEDARPNWPQVSLQYDSSGQWRHTCGGTLVDQSWLVTAACHICSSRTYRVVLGRH 88
QY 58 DRSSNAEPQVLSIARATIHNNANTMN--NDLTLLKLASPARVTAQVSPVCLASTNEA 115
DB :|||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:|
DB 89 SLSTNEPGSLAVKVKLVVHQDNGNDIALLKLASPSVSLTDKQLQCLPAAGTY 148
QY 116 LPSGLTCVTGMRISGVGNVTPARLQVYVPLVTVNQCRQ--YMGARITDAMICAGSG 173
DB :|||:| |||||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:|
DB 149 LPNNVVCYVTGWRGLQTNG-ASPDILQOQLLVVDYATCSKPGWGSTVKTNNICAGGDS 207
QY 174 -ASSCGDSGGPLVCQKGN-TWVLIGIVSWGTK-NCN-IOAPAMYTRVSKFSTWVNOVA 229
DB :|||:| ||||| |||:| |||:| |||:| |||:| |||:| |||:| |||:|
DB 208 IISSCNGDSGGPLNCQGANGOMQVHGVSFGSLGCGNYHKPSVETRVSNYIDWINSVIA 267
QY 230 YN 231
DB 268 NN 269

RESULT 10
EL2_RAT
ID EL2_RAT STANDARD; PRT; 271 AA.
AC P00774;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Elastase 2 precursor (EC 3.4.21.71).
GN ELA2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82182967; PubMed=6918221;
RA McDonald R.J., Swift G.H., Quinto C., Swain W., Pictet R.L.,
RA Nikovits W., Rutter W.J.;
RT "Primary structure of two distinct rat pancreatic preproelastases
RT determined by sequence analysis of the complete cloned messenger
RT ribonucleic acid sequences.";
RL Biochemistry 21:1453-1463(1982).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85054882; PubMed=6094548;
RA Swift G.H., Craik C.S., Stary S.J., Quinto C., Lahale R.G.,
RA Rutter W.J., Macdonald R.J.;
RT "Structure of the two related elastase genes expressed in the rat
RT pancreas.";
RL J. Biol. Chem. 259:14271-14278(1984).
CC -|- FUNCTION: ACTS UPON ELASTIN.
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Leu-|-Xaa, Met-|-Xaa
CC and Phe-|-Xaa. Hydrolyzes elastin.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- TISSUE SPECIFICITY: PANCREAS.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V01233; CAA24543.1; -.
DR EMBL; L00124; AAA98780.1; -.
DR EMBL; L00118; AAA98780.1; JOINED.
```

DR EMBL; L00119; AAA98780.1; JOINED.
DR EMBL; L00120; AAA98780.1; JOINED.
DR EMBL; L00121; AAA98780.1; JOINED.
DR EMBL; L00122; AAA98780.1; JOINED.
DR EMBL; L00123; AAA98780.1; JOINED.
DR PIR; A00961; ELRT2.
DR HSSP; P00772; IELG.
DR MEROPS; S01155; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydrolase; Serine protease; Zymogen; Signal.
FT SIGNAL 1 16
FT PROPEP 17 30
FT CHAIN 31 271
FT ACT_SITE 75 75
FT ACT_SITE 123 123
FT ACT_SITE 218 218
FT ACT_SITE 218 218
FT DISULFID 60 76
FT DISULFID 157 224
FT DISULFID 188 204
FT DISULFID 214 245
FT DISULFID 214 245
SQ SEQUENCE 271 AA; 28885 MW; 125C783B857B71E3 CRC64;

Query Match 39.9%; Score 499; DB 1; Length 271;
Best Local Similarity 43.0%; Pred. No. 1.2e-37;
Matches 104; Conservative 37; Mismatches 89; Indels 12; Gaps 8;

QY 1 IVNGENAVPGSMWQVSLQ---DNTGFHFCGSLISPNWVTAACHQVTPGRHFVVLGEY 57
DB 31 VVGQASPNWQVSLQYLSGKWHHTCGSLVANNVLAACHISRYRYVLLGRH 90
QY 58 DRSSNAEPQVLSIARAIHPNNANTM--NDLTLLKLSAPRYTAQVSPVCLASTNEA 115
DB 91 SLSTSESGSLAVQVSLVHVKWNAQKLSNGNDIALVLPALVLTQACLPAGTI 150
QY 116 LPGLTCVTGTGRISGVGNVTPARLQVPLVTVNQC--ROYWGARITDMICAGSG 173
DB 151 LPNNYPCYVTGRLQTNG--ATPDVLQOGRLLVVDYATCSSAWGSSVTKTNMVCAGDG 209
QY 174 -ASSCGDGGPLVCQKGN--TWVLIGVSWG-TKNCNI-QAPAMYTRVSKFSTWVNOVMA 229
DB 210 VTSSCNGDGGPLNCAQNSGQVHGVISFGSTLGCNYPKFSVFTRVSNYIDWINSVIA 269

QY 230 YN 231
DB 270 KN 271

RESULT 11
CAC3_BOVIN STANDARD; PRT; 253 AA.
ID CAC3_BOVIN
AC P05805;
DT 01-NOV-1988 (Rel. 09, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Proprietary E precursor (Procarboxypeptidase A complex component
DE III) (Procarboxypeptidase A-S6 subunit III) (PROCPA-S6 III).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE OF 1-25.
RX MEDLINE=91099520; PubMed=2269366;
RA Pascual R., Vendrell J., Aviles F.X., Bonicel J., Wicker C.,
RA Puigserver A.;

RT "Autolysis of proproteinase E in bovine procarboxypeptidase A ternary
RT complex gives rise to subunit III.";
RL FEBS Lett. 277:37-41(1990).
RN [2]
RP SEQUENCE OF 14-253, AND DISULFIDE BONDS.
RX MEDLINE=86220198; PubMed=3519215;
RA Venot N., Sciaky M., Puigserver A., Desnuelle P., Laurent G.;
RT "Amino acid sequence and disulfide bridges of subunit III, a
RT defective endopeptidase present in the bovine pancreatic 6 S
RT procarboxypeptidase A complex.";
RL Eur. J. Biochem. 157:91-99(1986).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
RX MEDLINE=94222022; PubMed=8168476;
RA Pignol D., Gaboriaud C., Michon T., Kerfelec B., Chapus C.,
RA Fontecilla-Camps J.C.;
RT "Crystal structure of bovine procarboxypeptidase A-S6 subunit III, a
RT highly structured truncated zymogen E.";
RL EMBO J. 13:1763-1771(1994).
CC -!- FUNCTION: DEFECTIVE ELASTASE-LIKE SERINE PROTEASE. DOES NOT SEEM
CC TO HAVE A PROTEASE ACTIVITY. ITS LIKELY FUNCTION IS TO PROTECT
CC PROCARBOXYPEPTIDASE A AGAINST DENATURATION IN THE ACIDIC
CC ENVIRONMENT OF THE RUMINANT DUODENUM.
CC -!- SUBUNIT: HETEROTRIMER OF SUBUNIT III; CARBOXYPEPTIDASE A AND
CC CHYMOTRYPSINOGEN C.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: PANCREAS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR PIR; A25065; CPBOA3.
DR PDB; 1FON; 14-OCT-96.
DR MEROPS; S01983; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Serine protease homolog; Pancreas; Digestion; 3D-structure.
KW Serine protease homolog; Pancreas; Digestion; 3D-structure.
FT PROPEP 1 11
FT CHAIN 12 253
FT DISULFID 41 57
FT DISULFID 100 103
FT DISULFID 140 206
FT DISULFID 171 187
FT DISULFID 196 227
SQ SEQUENCE 253 AA; 27337 MW; 24663724D8AE409C CRC64;

Query Match 39.7%; Score 497.5; DB 1; Length 253;
Best Local Similarity 43.6%; Pred. No. 1.5e-37;
Matches 106; Conservative 36; Mismatches 88; Indels 13; Gaps 8;

QY 1 IVNGENAVPGSMWQVSLQ---DNTGFHFCGSLISPNWVTAACHQVTPGRHFVVLGEY 57
DB 12 VVGQASPNWQVSLQYLSGKWHHTCGSLVANNVLAACHISRYRYVLLGRH 71
QY 58 DRSSNAEPQVLSIARAIHPNNANTM--NDLTLLKLSAPRYTAQVSPVCLASTN 113
DB 72 DRSVLEGESEQVPIINAGDLFVHPVHNSNCVACGNDIALVLPALVLTQACLPAG 131
QY 114 EALPSGLTCVTGTGRISGVGNVTPARLQVPLVTVNQCRO--TWGARITDMICAGG 171
DB 132 DILNEAPCYISGWGRLY-TGGPLPKLOQLLPVVDYEHCSQMDWMTIVTKTMVCA 190
QY 172 SGASSCGDGGPLVCQKGN--TWVLIGVSWG-TKNCNI-QAPAMYTRVSKFSTWVNOV 228
DB 191 DTRSGCNGDGGPLNCAQNSGQVHGVISFGSTLGCNYPKFSVFTRVSNYIDWIDETI 250

QY 229 YN 231
DB 251 ASN 253

DR	PRINTS; PRO0722; CHYMOTRYPSIN.
DR	SMART; SM00020; Tryp_SPC; 1.
DR	PROSITE; PS0240; TRYPSIN_DOM; 1.
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.
DR	PROSITE; PS00135; TRYPSIN_SER; 1.
KW	Hydrolase; Serine protease; Zymogen; Signal.
FT	SIGNAL 1 16
FT	PROPEP 17 28 ACTIVATION PEPTIDE.
FT	CHAIN 29 269 ELASTASE 2A.
FT	DISULFID 58 74 BY SIMILARITY.
FT	DISULFID 155 222 BY SIMILARITY.
FT	DISULFID 186 202 BY SIMILARITY.
FT	DISULFID 212 243 BY SIMILARITY.
FT	ACT_SITE 73 73 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 121 121 ACT SITE 121 121 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 216 216 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	CONFLICT 202 202 C -> V (IN REF. 3).
FO	SEQUENCE 269 AA; 28888 MW; A2E05143EF4987C CRC64;
 Query Match 38.8%; Score 486; DB 1; Length 269; Best Local Similarity 43.0%; Pred. No. 1.7e-36; Matches 104; Conservative 41; Mismatches 85; Indels 12; Gaps	
Qy	1 IVNGENAVPGSWPWQVSLQDNTG---FHFCGGSLIPNWWVTAAHCOVTPGRHFVVLEY 57 : : : : : : : : :
Db	29 VVGGEARPNSWPQWSLQYSSNGKWKYHTCGGLSIANSWVLTAHCITSSRTYRVGLGRH 88 : : : : : : : :
Qy	58 DRSSNAEPQVLSSTARITHPNNNANTMN--NDLTLLKLSPARYTAQVSPVCLASTNEA 115 : : : : : : : : : :
Db	89 NLYVAESGLASVSVKIVVHKDNWSNISKGNDIALLKLANVSLTDKIQLACLPPAGTI 148 : : : : : : : :
Qy	116 LPSCGLTCVTGWGRISGVGNVTFARLQQVPLPVTVNQCRQ--YWGARTIDAMICAGSG 173 : : : : : : :
Db	149 LPNNYPCTVGTGWLQTNGAV-PDVLOOGRLLVVDYATCSSSAWGSVKTSMICAGDG 207 : : : : : : :
Qy	174 A-SSCGDGSGPLVCQKGN-TWVLIGTIVSGTK-NCN-IOAPAMYTRVSKFSTWINOVMA 229 : : : : : : : :
Db	208 VISSCNGDSGPLNCQASDRGWQHVIGSVFSGRLGCNYHKPSVFTRVSNYIDWINSVIA 257 : : : : : : : :
Qy	230 YN 231
Db	268 NN 269
 RESULT 13 EL3B_HUMAN STANDARD; PRT; 270 AA. ID EL3B_HUMAN AC P08861; P11423; DT 01-NOV-1988 (Rel. 09, Created) DT 16-OCT-2001 (Rel. 40, Last sequence update) DT 15-JUN-2002 (Rel. 41, Last annotation update) DE Elastase IIIB precursor (EC 3.4.21.70) (Protease E). GN ELA3B. OS Homo sapiens (Human). OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. OX NCBI_TaxID=9606; RN [1] RP SEQUENCE FROM N.A. RC TISSUE=Pancreas; RX MEDLINE=88087253; PubMed=2826474; RA Tani T., Ohsumi J., Mita K., Takiguchi Y.; RT "Identification of a novel class of elastase isozyme, human pancreatic elastase III, by cDNA and genomic gene cloning."; RL J. Biol. Chem. 263:1231-1239(1988). RN [2] RP SEQUENCE FROM N.A. RC TISSUE=Pancreas; RA Strausberg R.; RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases. RN [3] RP SEQUENCE OF 4-270 FROM N.A. RC TISSUE=Pancreas;	

RX MEDLINE=88000545; PubMed=3477287;
RA Shen W., Fletcher T.S., Largman C.;
RT "Primary structure of human pancreatic protease E determined by
RT sequence analysis of the cloned mRNA.";
RL Biochemistry 26:3447-3452(1987).
RN [4]
RC SEQUENCE OF 31-50.
RP TISSUE=Pancreas;
RX MEDLINE=89325560; PubMed=2753124;
RA Moulard M., Kerfelec B., Mallet B., Chapus C.;
RT "Identification of a procarboxypeptidase A-truncated protease E
RT binary complex in human pancreatic juice.";
RL FEBS Lett. 250:186-170(1989).
RN [5]
RP SEQUENCE OF 94-164, AND CARBOHYDRATE-LINKAGE SITES ASN-114.
RX TISSUE=Pancreas;
RC MEDLINE=89289396; PubMed=2737288;
RA Wendorf P., Geyer R., Szegoleit A., Linder D.;
RT "Localization and characterization of the glycosylation site of human
RT pancreatic elastase 1.";
RL FEBS Lett. 249:275-278(1989).
RN [5]
CC -!- FUNCTION: EFFICIENT PROTEASE WITH ALANINE SPECIFICITY BUT ONLY
CC LITTLE ELASTOLYTIC ACTIVITY.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Ala-Xaa. Does not
CC hydrolyse elastin.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC -!- CAUTION: Was originally (Ref.5) thought to be elastase 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M16630; AAA36482.1; -
DR EMBL; BC005216; AAA05216.1; -
DR EMBL; M18692; AAA58454.1; -
DR PIR; B29934; B29934.
DR PIR; A27206; A27206.
DR PIR; S04999; S04999.
DR PIR; S04490; S04490.
DR HSP; P05805; IFON.
DR MEROPS; S01.205; -
DR GlycoSuiteDB; P08861; -
DR SWISS-2DPAGE; P08861; HUMAN.
DR Genew; HGNC:15945; ELA3B.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Zymogen; Signal; Glycoprotein.
FT SIGNAL 1 15 OR 16 (POTENTIAL).
FT PROPEP 16 28 ACTIVATION PEPTIDE (POTENTIAL).
FT CHAIN 29 270 ELASTASE IIIB.
FT ACT_SITE 73 73 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 123 123 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 217 217 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 58 74 BY SIMILARITY.
FT DISULFID 117 120 PROBABLE.
FT DISULFID 157 223 BY SIMILARITY.
FT DISULFID 188 204 BY SIMILARITY.
FT DISULFID 213 244 BY SIMILARITY.
FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .).
FT FTID-CAR_000212.
FT CONFLICT 4 4 R -> G (IN REF. 3).
FT CONFLICT 64 64 A -> G (IN REF. 3).
FT CONFLICT 79 79 W -> R (IN REF. 1).

FT CONFLICT 129 131 MISSING (IN REF. 5).
FT CONFLICT 164 164 R -> P (IN REF. 3).
SQ SEQUENCE 270 AA; 29293 MW; B14BE0AAD369SAFE CRC64;
Query Match 38.5%; Score 481.5; DB 1; Length 270;
Best Local Similarity 42.7%; Pred. No. 4.3e-36;
Matches 103; Conservative 39; Mismatches 86; Indels 13; Gaps 9;
QY 1 IVNGENAVPGSPWQVSLQ-DNTG--FHPGGSLSIPNVTAAHCOVTPGRHVFVLGY 57
DB 29 VVNGEDAVPYSPWQVSLQYKSGSFYHTCGSLAPDWVVTAGHCISSTWYQVVLGY 88
QY 58 DRSSNAEPVQVLSI--ARAITHPNMNATM--NNDTLTLKLASPARYTAQVSPVCLASTN 113
DB 89 DRAVKEGPEQVTPINSGLDFVHPLNRSVACGNDIALIKLSAQLGDAVQLASLPPAG 148
QY 114 EALPSGLICVTGWRISGVGNVTPARLQOVVLPVTVNQCRQ--YWGARIYDAMTCAGG 171
DB 149 DILPNETPCYITGWRGLYTNGLPL-PDKLOEALLPVVDYEHCSRWNWMSGVKTKMYCAGG 207
QY 172 SGASSCOGDSGGPLVC-OKGNTWVLIGIVSW-GTKNCNI-QAPAMVTRYSKFSTWNOVM 228
DB 208 DIRSCNGDSGGPLNCPTEDEGGQVHGVSFVSAGFCNTRRPTVTRVSADFIDWIEETI 267
QY 229 A 229
DB 268 A 268
RESULT 14
EL2_BOVIN STANDARD; PRT; 269 AA.
AC Q29461;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Elastase 2 precursor (EC 3.4.21.71).
GN ELA2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=98079203; PubMed=9418008;
RA Gestin M., le Huerou-Luron I., Wicker-Planquart C., le Drean G.,
RA Chaix J.C., Puigserver A., Guilloteau P.;
RT "Bovine pancreatic preproelastases I and II: comparison of nucleotide
RT and amino acid sequences and tissue specific expression.";
RL Comp. Biochem. Physiol. 118B:181-187(1997).
CC -!- FUNCTION: ACTS UPON ELASTIN.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Leu-Xaa, Met-Xaa
CC and Phe-Xaa. Hydrolyzes elastin.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: PANCREAS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X97635; CAA66231.1; -
DR HSP; P00766; ICHG.
DR MEROPS; S01.155; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.


```
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase: Serine protease; Zymogen; Signal.
FT SIGNAL 1 16
FT PROPEP 17 269
FT CHAIN 29 269
FT ACT_SITE 73 73
FT ACT_SITE 121 121
FT ACT_SITE 216 216
FT DISULFID 58 74
FT DISULFID 155 222
FT DISULFID 186 202
FT DISULFID 212 243
SQ SEQUENCE 269 AA; 28856 MW; 8343897062CF267C CRC64;

Query Match 38.3%; Score 480; DB 1; Length 269;
Best Local Similarity 42.6%; Pred. No. 5.9e-36;
Matches 103; Conservative 39; Mismatches 88; Indels 12; Gaps 8;

Qy 1 IVNGENAVPGSWPQVSLQDNTG---FHFCGSLSPNNVYTAHCQVTPGRHFVVLGEY 57
Db 29 VVGDEDARPNPQVSLQYSSGOWRHTCGSLIEQNWVLTAAHCISSSRTYRVVGRQ 88
Qy 58 DRSSNAEPVQVLSIARATHPNWNTNM--NNDTLKLLASPARVTAQVSPVCLASTNEA 115
Db 89 SLSTVESGSLTIAYSKSVIHEKWSNQLAQNDAALLKASSVPLTDKIQLGCLPAAGTI 148
Qy 116 LPSSLTCTVTTGWRISGVGNVTPARLQVVLPLVTNQC--QYWGARITDAMICAGSG 173
Db 149 LPNNYVCVYTGWRISGVGNVTPARLQVVLPLVTNQC--QYWGARITDAMICAGSG 207
Qy 174 -ASSCGDGGPLVCQKGN-TWVLIGIVSWGK-NCN-IOAPAMYTRVSKFSTWLNQVMA 229
Db 208 VTSSCGDGGPLNCQAANRQWQVHGVISFGSLGCGNYRKPVSFVTRVSNNDWISSVIE 267
Qy 230 YN 231
Db 268 NN 269

RESULT 15
EL2_MOUSE STANDARD; PRT; 271 AA.
AC P05208;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DE Elastase 2 precursor (EC 3.4.21.71).
GN ELA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-87066713; PubMed-3641189;
RT Stevenson B.J., Hagenbuechle O., Wellauer P.K.;
RT "Sequence organisation and transcriptional regulation of the mouse
RT elastase II and trypsin genes.";
RL Nucleic Acids Res. 14:8307-8330(1986).
CC -1- FUNCTION: ACTS UPON ELASTIN.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Leu-|-Xaa, Met-|-Xaa
CC and Phe-|-Xaa. Hydrolyzes elastin.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: PANCREAS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04573; CAA28242.1; -.
DR PIR; A25528; A25528.
DR HSSP; P00772; IELG.
DR MEROPS; S01155; -.
DR MGD; MGI:95316; Ela2.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase: Serine protease; Zymogen; Signal.
FT SIGNAL 1 16
FT PROPEP 17 30
FT CHAIN 31 271
FT ACT_SITE 75 75
FT ACT_SITE 123 123
FT ACT_SITE 218 218
FT DISULFID 60 76
FT DISULFID 157 224
FT DISULFID 188 204
FT DISULFID 214 245
SQ SEQUENCE 271 AA; 28913 MW; FA542AE38FD3B4B CRC64;

Query Match 37.9%; Score 475; DB 1; Length 271;
Best Local Similarity 40.5%; Pred. No. 1.7e-35;
Matches 98; Conservative 39; Mismatches 93; Indels 12; Gaps 8;

Qy 1 IVNGENAVPGSWPQVSLQ---DNTGFHFCGSLSPNNVYTAHCQVTPGRHFVVLGEY 57
Db 31 VVGQEAETPTWQVSLQVLSGSRWRHNCGGSLVANNVLTAAHCLSNYQTYRVLLGAH 90
Qy 58 DRSSNAEPVQVLSIARATHPNWNTNM--DLTLLKLLASPARVTAQVSPVCLASTNEA 115
Db 91 SLSPGAGSAAVQVSKLVVHORWNSQNGVDIALIKLASPVTLKNIQTACLPPAGTI 150
Qy 116 LPSSLTCTVTTGWRISGVGNVTPARLQVVLPLVTNQC--QYWGARITDAMICAGSG 173
Db 151 LPNNYVCVYTGWRISGVGNVTPARLQVVLPLVTNQC--QYWGARITDAMICAGSG 209
Qy 174 -ASSCGDGGPLVCQKGN-TWVLIGIVSWGK-NCN-IOAPAMYTRVSKFSTWLNQVMA 229
Db 210 VTSSCGDGGPLNCRAASNGQWQVHGVISFGSLGCGNYRKPVSFVTRVSNYIDWINSVMA 269
Qy 230 YN 231
Db 270 RN 271
```

Search completed: December 20, 2002, 15:17:15
Job time : 8.91304 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:14:28 ; Search time 14.6087 Seconds
(without alignments)
1520.126 Million cell updates/sec

Title: US-09-856-319B-4_COPY_34_264

Perfect score: 1252

Sequence: 1 IVNGENAVPGSWPQVSLQD.....AMTVRSKFTWVNOVMAYN 231

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73.*

1: p1r1.*

2: p1r2.*

3: p1r3.*

4: p1r4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1119	89.4	264	2	I38136 chymotrypsin-like
2	757.5	60.5	263	2	A21195 chymotrypsin (EC 3
3	750.5	59.9	263	2	S47537 chymotrypsin (EC 3
4	750.5	59.9	263	2	A31299 chymotrypsin (EC 3
5	740.5	59.1	245	1	KYBOB chymotrypsin (EC 3
6	727.5	58.1	245	1	KYBOA chymotrypsin (EC 3
7	715.5	57.1	263	1	KVRTB chymotrypsin (EC 3
8	714.5	57.1	244	2	S72219 chymotrypsin B - A
9	503	40.2	269	2	A26823 pancreatic elastase
10	499	39.9	271	1	ELRT2 pancreatic elastase
11	487	38.9	1524	2	T30337 polyprotein - Afri
12	486	38.8	269	2	B26823 pancreatic elastase
13	485.5	38.8	240	1	CPBOA3 procarboxypeptidase
14	481.5	38.5	270	2	B29934 pancreatic elastase
15	475	37.9	271	2	A25528 pancreatic elastase
16	464.5	37.1	810	1	PLHU plasmin (EC 3.4.21
17	462	36.9	812	1	PLMS plasmin (EC 3.4.21
18	460.5	36.8	270	2	A29934 pancreatic elastase
19	455.5	36.4	270	2	S56160 mast cell tryptase
20	452.5	36.1	460	2	B61545 plasmin (EC 3.4.21
21	451.5	36.1	810	2	B30848 plasmin (EC 3.4.21
22	449.5	35.9	790	1	PLPG plasmin (EC 3.4.21
23	449	35.9	269	2	C26823 pancreatic elastase
24	448.5	35.8	343	1	A57014 prostasin (EC 3.4.
25	447.5	35.7	268	2	S68826 pancreatic elastase
26	447.5	35.7	268	2	S68826 pancreatic elastase
27	447.5	35.7	273	2	A47246 trypsin (EC 3.4.2
28	444	35.5	231	1	TRPGTR trypsin (EC 3.4.21
29	441	35.2	126	2	A23473 chymotrypsin-like

30 440 35.1 246 1 TRRT1 trypsin (EC 3.4.21

31 439.5 35.1 274 2 JC4171 trypsin (EC 3.4.2

32 437.5 34.9 455 2 A61545 plasmin (EC 3.4.21

33 436 34.8 246 1 TRRT2 trypsin (EC 3.4.21

34 436 34.6 246 1 TRDGC trypsin (EC 3.4.21

35 433 34.6 247 1 TRDGC trypsin (EC 3.4.21

36 432 34.5 248 2 S55067 trypsin (EC 3.4.21

37 430.5 34.4 266 1 ELRT1 pancreatic elastase

38 430.5 34.4 276 2 A38654 mast cell proteina

39 430 34.3 238 2 S31779 trypsin (EC 3.4.21

40 429 34.3 1019 1 A56318 enteropeptidase (E

41 428.5 34.2 275 2 A32410 trypsin (EC 3.4.2

42 428 34.2 810 2 I46260 plasmin (EC 3.4.21

43 427.5 34.1 258 2 S70439 pancreatic elastase

44 427.5 34.1 267 4 A56613 probable pancreati

45 427.5 34.1 275 2 C35863 trypsin (EC 3.4.2

ALIGNMENTS

RESULT 1

I38136
chymotrypsin-like proteinase (EC 3.4.21.-) CTRL-1 - human
C:Species: Homo sapiens (man)
C:Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 22-Jun-1999
C:Accession: I38136
R:Larsen, F.; Solheim, J.; Kristensen, T.; Kolsto, A.B.; Prydz, H.
Hum. Mol. Genet. 2, 1589-1595, 1993
A:Title: A tight cluster of five unrelated human genes on chromosome 16q22.1.
A:Reference number: I38135; MUID:94093544; PMID:8268911
A:Accession: I38136
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-264 <RES>
A:Cross-references: EMBL:X71874; NID:g406226; PIDN:CAA50710.1; PID:g406228
C:Genetics:
A:Gene: GDB:CTRL
A:Cross-references: GDB:204061
A:Map position: 16q22.1-16q22.1
A:Introns: 18/1; 52/3; 79/2; 106/3; 167/1; 211/3
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:34-257/Domain: trypsin homology <TRY>
F:75,121,214/Active site: His, Asp, Ser #status predicted

Query Match 89.4%; Score 1119; DB 2; Length 264;
Best Local Similarity 86.1%; Pred. No. 6.2e-90;
Matches 199; Conservative 22; Mismatches 10; Indels 0; Gaps 0;

QY 1 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVTAACHTVTPGRHFVVLGYDRS 60

Db 34 IVNGENAVLGSWPMQVSLQDSSGFHFCGSLISQSWVTAACHTVTPGRHFVVLGYDRS 93

QY 61 SNAEPQVLSIARATHPNWNTANNNDTLTKLSPARYTAQVSPVCLASTNEALPSGL 120

Db 94 SNAEPQLVLSVSRATHPSWNTTNDVTLTKLSPAQYTTTRISPVCLASSNEALTEGL 153

QY 121 TCVTTGWGRISGVNVTAPLQOVPLVTVNCQRCYWGARTDAMICAGGSSCCQGD 180

Db 154 TCVTTGWGRISGVNVTAPLQOVPLVTVNCQRCYWGSSITDSMICAGGSSCCQGD 213

QY 181 SGGPLVCOKGNTWLLIGIVSWGTKNCNICAPAMTVRSKFTWVNOVMAYN 231

Db 214 SGGPLVCQKNTWLLIGIVSWGTKNCNRAPAVTVTRSKFTWVNOVIAYN 264

RESULT 2

A21195

chymotrypsin (EC 3.4.21.1) 2 precursor - dog

C:Species: Canis lupus familiaris (dog)

C:Date: 27-Jul-1990 #sequence_revision 27-Jul-1990 #text_change 22-Jun-1999

C:Accession: A21195

R;Pinsky, S.D.: LaForge, K.S.: Luc, V.: Scheele, G.
Proc. Natl. Acad. Sci. U.S.A. 80, 7486-7490, 1983
A;Title: Identification of cDNA clones encoding secretory isoenzyme forms: sequence deter-

A;Reference number: A21195; MUID:84170253; PMID:6584866
A;Accession: A21195
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-263 <PIN>
A:Cross-references: GB:K01173; NID:g163945; PIDN:AAA30841.1; PID:g163946
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; protein digestion; serine proteinase
F:34-256/Domain: trypsin homology <TRY>
F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 60.5%; Score 757.5; DB 2; Length 263;
Best Local Similarity 58.0%; Pred. No. 1.7e-58;
Matches 134; Conservative 35; Mismatches 61; Indels 1; Gaps 1;

Qy 1 IVNGENAVPGSWPQVSLDNTGFHFCGSSLSPNNWVTAAHCQVTPGRHFVVLGEYDRS 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 34 IVNGEADVPGSWPQVSLDSTGFHFCGSSLSEDWVVTAACGVRT-THQVVAEPDQG 92
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 61 SNAEPVQVLSTARATHPNWNANTMNNDLTLLKLASPARYAQAQSPVCLASTNEALPSGL 120
||:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db 93 SDAESIQVLKIATFKAKNPKNFTINDDITLKLATPAKFSTVSAYCLPQATDDFPAGT 152
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 121 TCVTTGWGRISGVNVTPARLOOVPLPTVYNQCROYWGARIITAMICAGSGASSCOGD 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 153 LCVTTGGWLTKHTNANTPDKLOOAAAPLLSNAECKFWGSKITDLMLWCAGASGVSSCMGD 212
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 181 SGGPLVCOKGNWTWLVIGVSWGKTNCNKTCQAPMAYTRVSKFTWINQVMAYN 231
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 213 SGGPLVCOKDGAWTLVGIVSWGSGTCTSTFGVYARVTKLIIPWQQIILQAN 263
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
S47537
chymotrypsin (EC 3.4.21.1) precursor - Atlantic cod
C:Species: Gadus morhua (Atlantic cod)
C:Date: 26-Dec-1994 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999
C:Accession: S47537; S43163
R:Gundmundsdottir, A.; Oskarsson, S.; Bakin, A.E.; Craik, C.S.: Bjarnason, J.B.
Biochim. Biophys. Acta 1219, 211-214, 1994
A>Title: Atlantic cod cDNA encoding a psychrophilic chymotrypsinogen.
A:Reference number: S47537; MUID:94368860; PMID:8086467
A:Accession: S47537
A:Molecule type: mRNA
A:Residues: 1-263 <GUD>
A:Cross-references: EMBL:X78490; NID:q468750; PIDN:CAA55242.1; PID:q468751
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; protein digestion; serine proteinase
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-263/Product: chymotrypsin #status predicted <MAT>
F:34-256/Domain: trypsin homology <TRY>
F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 59.9%; Score 750.5; DB 2; Length 263;
Best Local Similarity 58.4%; Pred. No. 6.9e-58;
Matches 135; Conservative 33; Mismatches 62; Indels 1; Gaps 1;

Qy 1 IVNGENAVPGSWPQVSLDNTGFHFCGSSLSPNNWVTAAHCQVTPGRHFVVLGEYDRS 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 34 IVNGEAVPHSWQVSLDQTFHFCGSSLINENWVVTAAHCNV-KNYHRVVLGEHDRS 92
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 61 SNAEPVQVLSTARATHPNWNANTMNNDLTLLKLASPARYAQAQSPVCLASTNEALPSGL 120
||:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db 93 SNSEGVQVMTVQVKHPRYNGFTINDDILLVKLATPATLNMRVSPVCLAETDDVFEGGM 152
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 121 TCVTTGWGRISGVNVTPARLOOVPLPTVYNQCROYWGARIITAMICAGSGASSCOGD 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 153 KCGTSGWGLTRYNAADPTALLQOAAAPLLTNPCQCKWGNKITSIDLMIAGAAGASCSMGD 212
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 181 SGGPLVCOKGNWTWLVIGVSWGKTNCNKTCQAPMAYTRVSKFTWINQVMAYN 231

F;16-238/Domain: trypsin homology <TRY>
F;1-122,42-58,136-201,168-182,191-220/Duplicate bonds: #status experimental
F;57,102,195/Active site: His, Asp, Ser #status experimental

	Query Match	58.1%	Score 727.5;	DB 1;	Length 245;
	Best Local Similarity	56.7%;	Pred. No. 6.4e-56;		
	Matches 131;	Conservative 32;	Mismatches 67;	Indels 1;	Gaps
Qy	1	IVNGENAVPGSWPQVWSLODNTGFHFCGGLSLIPNNVYVTAACHCQVTPGRHFFVVLGEYDRS	60		
Db	16	IVNGEAVPGSWPQVWSLODKTGFHFCGGLSLINNNVYVTAACHGVTTTS-DVVVAGEFPDQ	74		
Qy	61	SNAPVQVLSIARIATHPNMNNANTMNDITLLKLASPARYTAQVSPVCLASTNEALPSGL	120		
Db	75	SSSEKIQKLIKAVFKNSKYNLSLTINDITLLKLSTAAQSFSQTSVAVCLPSASDDFAAGT	134		
Qy	121	TCVTTGGRISGVGNVTPARLQOVVLPVTVNOCROYGCARITDAMICAGGSGASSCCGD	180		
Db	135	TCVTTGMGLTRYTNANTPORKQASLPLSNTNCKYGTGKTKIDAMICAGASGVSSCCMGD	194		
Qy	181	SGGPLVCQKQNTWVLIGIVSWGTFKNCNIQAPAMTRYRVSKFSTWQVYNN	231		
Db	195	SGGPLVCKKNGATLVIGIVSWGSSCTSTPGVIYARVYALVNNVQQTAAAN	245		

RESULT 7
KYRTB
chymotrypsin (EC 3.4.21.1) B precursor - rat
N/Alternate names: chymotrypsinogen B
C/Species: Rattus norvegicus (Norway rat)
C/Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 18-Jun-1999
C/Accession: A22658
J. Bell, G.I.; Quinto, C.; Quiroga, M.; Valenzuela, P.; Craik, C.S.; Rutter, J. Biol. Chem. 259, 14265-14270, 1984
A>Title: Isolation and sequence of a rat chymotrypsin B gene.
A/Reference number: A22658; MUID:85054881; PMID:6209274
A/Accession: A22658
A/Molecule type: DNA
A/Residues: 1-263 <BEL>
A/Cross-references: GB:K02298; NID:g203653; PIDN:AAA98732.1; PID:g203654
C/Genetics:
A/Introns: 18/1; 52/3; 79/2; 105/3; 166/1; 210/3
C/Superfamily: trypsin; trypsin homology
C/Keywords: hydrolase; pancreas; protein digestion; serine proteinase
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-33/Domain: propeptide #status predicted <PRO>
F;34-263/Product: chymotrypsin B #status predicted <MAT>
F;34-236/Domain: trypsin homology <TRY>
F;75,120,213/Active site: His, Asp, Ser #status predicted

	Query Match	57.1%	Score 715.5;	DB 1;	Length 263;
	Best Local Similarity	55.4%;	Pred. No. 7.6e-55;		
	Matches 128;	Conservative 35;	Mismatches 67;	Indels 1;	Gaps
Qy	1	IVNGENAVPGSWPQVWSLODNTGFHFCGGLSLIPNNVYVTAACHCQVTPGRHFFVVLGEYDRS	60		
Db	34	IVNGEADIPGSPWPQVWSLODKTGFHFCGGLSLIEDWVWVTAACHGVKTS-DVVVAGEFPDQ	92		
Qy	61	SNAPVQVLSIARIATHPNMNNANTMNDITLLKLASPARYTAQVSPVCLASTNEALPSGL	120		
Db	93	SDEENIQVLIQAQVFNKPFNMFTVRNDITLLKLATPAQFSQTSVAVCLPNVYDDDFPPT	152		
Qy	121	TCVTTGGRISGVGNVTPARLQOVVLPVTVNOCROYGCARITDAMICAGGSGASSCCGD	180		
Db	153	VCATTGMGKYKYNALTKPEKLOQAALPIVSEADCKKSWGSKITDVTTCAGASGVSSCCMGD	212		
Qy	181	SGGPLVCQKQNTWVLIGIVSWGTFKNCNIQAPAMTRYRVSKFSTWQVYNN	231		
Db	213	SGGPLVCQKQGVVTLAGIVSGVSGVCSSTPPIYVSRVYALPMPVQVQILEAN	263		

RESULT 8
S72219

Query Match 38.9%; Score 487; DB 2; Length 1524;
Best Local Similarity 39.5%; Pred. No. 4.2e-34;
Matches 94; Conservative 49; Mismatches 75; Indels 20; Gaps 6;

Qy 1 IVNGENAVPGSWPQVSLQDNTG---FHFCGSLISPNWVYTAHCOVTPGRHFVVLGEYDRS 60
Db 1295 VVGQQAAPRSWPLVLSIQNSKKRHYCGGIIITNKWILTAAHCEVKINLHRVYVGHDTLT 1354
Qy 61 --SNAEPQVLSIARATHPNWNTMNDLTLLKLASPARYTAQVSPVCLASTNEALPS 118
Db 1355 EVQNEHAKYVNSHVELYMP--GSPPRNDLLLELDTPLLLNSVAVIC-----LPD 1405
Qy 119 GLT-----CVTTGWRISGVNVT--PARLQVVLPLVTVNCRQYWGARTDAMICAG 170
Db 1406 DVTDTQAECLVAGWG-VTDVGMSLPTKLQAKVPIVSTFKCKDYVSDVTDNNICAG 1464
Qy 171 GSGASSCGDGGPLVCOKGNWTVLIGIVSWGTRKNCNTQAPAMVTRVSKFTWVNOVM 228
Db 1465 KAGASSCMGDSGGLICKREDRYVLGVVSWGSGKCDVKAPSVYTLTSAFMDWISQHM 1522

RESULT 12
B26823
pancreatic elastase II (EC 3.4.21.71) A precursor - human
C:Species: Homo sapiens (man)
C:Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 20-Jun-2000
C:Accession: B26823; A27432; A41431; S34491
R:Kawashima, I.; Tani, T.; Shimoda, K.; Takiguchi, Y.
DNA 6, 163-172, 1987
A:Title: Characterization of pancreatic elastase II cDNAs: two elastase II mRNAs are expressed in the pancreas of the rat
A:Reference number: A90958; MUID:87217962; PMID:3646943
A:Accession: B26823
A:Molecule type: mRNA
A:Residues: 1-269 <KAW>
A:Cross-references: GB:M16652; NID:g182057; PIDN:AAA52380.1; PID:g182058
R:Fletcher, T.S.; Shen, W.F.; Largman, C.
Biochemistry 26, 7256-7261, 1987
A:Title: Primary structure of human pancreatic elastase 2 determined by sequence analysis
A:Reference number: A27432; MUID:88107669; PMID:3427074
A:Accession: A27432
A:Molecule type: mRNA
A:Residues: 1-269 <FLE>
A:Cross-references: GB:M16631; NID:g182022; PIDN:AAA52374.1; PID:g182023
R:Shiraga, Y.; Yoshida, H.; Matsuki, S.; Takemura, K.; Ikeda, N.; Shimada, Y.; Ozawa, T.
J. Biochem. 102, 1553-1563, 1987
A:Title: Molecular cloning and expression in Escherichia coli of a cDNA encoding human pancreatic elastase II
A:Reference number: A41431; MUID:88198076; PMID:2834346
A:Accession: A41431
A:Molecule type: mRNA
A:Residues: 1-201, 'V', 203-269 <SHI>
A:Cross-references: GB:D00236; NID:g219619; PIDN:BAA00165.1; PID:g219620
A:Note: the authors translated the codon GTG for residue 202 as Cys
R:Moulard, M.; Michon, T.; Kerfelec, B.; Chapus, C.
FEBS Lett. 261, 179-183, 1990
A:Title: Further studies on the human pancreatic binary complexes involving procarboxypeptidase A
A:Reference number: S08253; MUID:90169111; PMID:2307232
A:Accession: S34491
A:Molecule type: protein
A:Residues: 'X', 18-50 <MOU>
C:Genetics:
A:Gene: GDB:ELAI
A:Cross-references: GDB:l19866; OMIM:l30120
A:Map position: 12pter-12qter
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; serine proteinase
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-28/Domain: propeptide #status predicted <PRO>
F:29-269/Product: pancreatic elastase IIA #status predicted <MAT>
F:29-262/Domain: trypsin homology <TRY>
F:73,121,216/Active site: His, Asp, Ser #status predicted

Query Match 38.8%; Score 486; DB 2; Length 269;
B29934

Best Local Similarity 43.0%; Pred. No. 7.1e-35;
Matches 104; Conservative 41; Mismatches 85; Indels 12; Gaps 8;

Qy 1 IVNGENAVPGSWPQVSLQDNTG---FHFCGSLISPNWVYTAHCOVTPGRHFVVLGEY 57
Db 29 VVGEEARPNSWPQVSLQYSGKWKYHTCGSLTANSWVLTAAHCTSSRTYRVGLGRH 88
Qy 58 DRSSNAEPQVLSIARATHPNWNTMNDLTLLKLASPARYTAQVSPVCLASTNEA 115
Db 89 NLYVAESGSLAVSVSKIVVHKDWSNQISKGNIDALLKLANPVLTDKIQIACLPAGTI 148
Qy 116 LPSGLTCVTTGWRISGVNVT--PARLQVVLPLVTVNCRQ--YWGARTDAMICAGSG 173
Db 149 LPNNYPCVVTGWRLOTNGAV--PDVLQOGRLLVVDYATCSSAWMGSSVKTSMICAGDG 207
Qy 174 A-SSCOGDSGGPLVCOKGN--TWVLIGIVSWGK--NCN--IQAPAMVTRVSKFTWVNOVM 229
Db 208 VISSCNGDSGGPLNCOASDGRMVHGIYFSGRLGCTNYHKPSVFTVRSVFNIDWINSVIA 267
Qy 230 YN 231
Db 268 NN 269

RESULT 13
CPBOA3
procarboxypeptidase A complex component III - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 18-Jul-1997
C:Accession: A25065
R:Venot, N.; Sciaky, M.; Puigserver, A.; Desnuelle, P.; Laurent, G.
Eur. J. Biochem. 157, 91-99, 1986
A:Title: Amino acid sequence and disulfide bridges of subunit III, a defective endopeptidase from the pancreas of the bovine
A:Reference number: A25065; MUID:86220198; PMID:3519215
A:Accession: A25065
A:Molecule type: protein
A:Residues: 1-240 <VBN>
C:Comment: This protein is found in the pancreatic juice of ruminants, where it constitutes a major component of the pancreatic secretory granules. It is inactive towards the usual substrates of the trypsin family: trypsin, chymotrypsin, and elastase. It is a defective endopeptidase against denaturation in the acidic environment of the ruminal duodenum.
C:Superfamily: trypsin; trypsin homology
C:Keywords: pancreas
F:1-233/Domain: trypsin homology <TRY>
F:28-44, 87-90, 127-193, 158-174, 183-214/Disulfide bonds: #status experimental
F:43,93,187/Active site: His, Asp, Ser #status predicted

Query Match 38.8%; Score 485.5; DB 1; Length 240;
Best Local Similarity 43.2%; Pred. No. 6.9e-35;
Matches 104; Conservative 36; Mismatches 88; Indels 13; Gaps 8;

Qy 3 NGENAVPGSWPQVSLQ---DNTGFHFCGSLISPNWVYTAHCOVTPGRHFVVLGEYDR 59
Db 1 DGEDAVPYVSWQVSLQYKEKGAFHHTCGSLIADPWVYVYTAGHCISTRTYQVVLGEYDR 60
Qy 60 S--SNAEPQVLSIARATHPNWNTMNDLTLLKLASPARYTAQVSPVCLASTNEA 115
Db 61 SVLEGEQVIPINAGDLFVHPLWNSVCACGNDIALVKLSRQAQLGDKVQLANLPAGDI 120
Qy 116 LPSGLTCVTTGWRISGVNVT--PARLQVVLPLVTVNCRQ--YWGARTDAMICAGSG 173
Db 121 LPNEAPCVISGGRLY--TGGPLPDKLQALLPVVDYEHCSQDWMWGTIVKTKTMVAGGT 179
Qy 174 ASSCOGDSGGPLVCOKGN--TWVLIGIVSW--GTKNCN--IQAPAMVTRVSKFTWVNOVM 230
Db 180 RSGCNGDSGGPLNCPADGSGQVHGVTSFVSAGCNTIKKPTVFTVRSVAFIDWIDETIAS 239
Qy 231 N 231
Db 240 N 240

RESULT 14
B29934

```

QY 58 DRSSNAEPVQLSI--ARAITHPNWANTM--NNDLTILKLASPARYTAQVSPVCLASTIN 111
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 89 DRAVKEGPEQIVT INSGDLFVHPHLMNRSCVACCDNALIKLSRAQLGDAVQLASLPPAG 148
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

QY 114 EALPSGLTCTVTTGWRISGVGNVTPARLQOVPLVTVNQCRQ--YWGARIIDAMICAGG 171
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 149 DILPNETPCYITGWRILYNGPL-PKQLQEAALLPVVDYEHCSRWNWWSGVKKTWVCAGG 207
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

QY 172 SGASSCGQSGGGLVLC-QKGNWTWVLIGIVSW-GTKNCNI-QAPAMYTRVSKFSTWINQVM 228
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 208 DIRSGCNGDSGGPLNCPTEPDGGQWGHGVTSFVSAGFCNTRRKPVTFTVRVSFAFIDMTEETI 267
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 229 A 229
   |
Db 268 A 268

RESULT 15
A25528
pancreatic elastase II (EC 3.4.21.71) precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 22-Jun-1999
C:Accession: A25528
R:Stevenson, B.J.; Hagenbuechle, O.; Wellauer, P.K.
Nucleic Acids Res. 14, 8307-8330, 1986
A:Title: Sequence organisation and transcriptional regulation of the mouse
A:Reference number: A93646; MUID:87066713; PMID:3641189
A:Accession: A25528
A:Molecule type: mRNA
A:Residues: 1-271 <STE>
A:Cross-references: GB:X04573; NID:g50825; PIDN:CAA28242.1; PID:g50826
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:1-30/Domain: signal sequence #status predicted <STG>
F:31-271/Product: pancreatic elastase II #status predicted <MAT>
F:31-264/Domain: trypsin homology <TRY>
F:75,123,218/Active site: His, Asp, Ser #status predicted

Query Match 37.9%; Score 475; DB 2; Length 271;
Best Local Similarity 40.5%; Pred.No. 6.5e-34;
Matches 98; Conservative 39; Mismatches 93; Indels 12; Gaps

QY 1 IVNGENAVPGSGWPQVWSLQ--DNTGFHCGGSLISPNVVTAAHCQVTPGRHFVVLGEV 57
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 31 VVGGEATPTNTWPQVSLQVLSSGRWRHNCGGLVANNVLTIAHCLSNKYQYRVLGGAH 90
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

QY 58 DRSSNAEPVQLSIARAITHPNWANTMNN--DLTLKLASPARYTAQVSPVCLASTNEA 115
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 91 SLSNPGAGSAAVQVSKLVVHORWNSQNVGNGYDIALIKLASPVTLTKNIQTACLPPAGTI 150
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

QY 116 LPSGLTCTVTTGWRISGVGNVTPARLQOVPLVTVNQCRQ--ROYWGARITDAMICAGGSG 173
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151 LPRNYCVYVGTGWLQTPNGN-SPDTLRQGRLLVVDYATCSSAWGSSVKKSSMVCAGGDG 209
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 174 -ASSCGDGGGLVLCOKGN-TWVLIGIVSGTKG-NCNI-QAPAMYTRVSKFSTWINQVMA 229
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 210 VTSSCNGDSGGPLNCRASNGQVGHVIFSGSLGNCNYPKPSVFTVRVSNYIDWINSVMA 269
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

QY 230 YN 231
   |
Db 270 RN 271

Search completed: December 20, 2002, 15:19:50
Job time : 15.6087 secs

```


11	481.5	38.5	242	7	AAP60061	Sequence of human
12	471.5	37.7	252	7	AAP60058	Sequence of human
13	471	37.6	279	22	ABG20513	Novel human diagno
14	469	37.5	269	7	AAP61723	Human elastase II.
15	465.5	37.2	791	18	AAW34285	Human plasminogen.
16	465.5	37.2	810	23	ABH83470	Human plasminogen.
17	465.5	37.2	810	23	ABH75939	Human plasminogen.
18	465.5	37.2	1220	22	AAU32858	Novel human secret
19	464.5	37.1	230	20	AAV31159	Human plasminogen
20	464.5	37.1	249	23	ABH81497	Human microplasmin
21	464.5	37.1	348	23	ABH81498	Human miniplasmino
22	464.5	37.1	437	19	AAW51457	Human plasminogen
23	464.5	37.1	546	14	AAK34427	Sequence of tissue
24	464.5	37.1	790	15	AAH60519	Human 'Glu' plasmi
25	464.5	37.1	790	22	AAH36562	Mammalian kringle
26	464.5	37.1	791	21	AAH01887	Human plasminogen.
27	464.5	37.1	791	21	AAV99589	Human plasminogen.
28	464.5	37.1	791	21	AAV50867	Human plasminogen
29	464.5	37.1	791	22	AAH67223	Amino acid sequenc
30	464.5	37.1	791	23	AAH23660	Human plasminogen
31	464.5	37.1	810	11	AAH08065	Human plasminogen
32	464.5	37.1	810	12	AAH31219	R561G human plasmi
33	464.5	37.1	810	12	AAH12406	R561E human plasmi
34	464.5	37.1	810	12	AAH13220	R561S human plasmi
35	464.5	37.1	810	12	AAH13221	Human plasminogen
36	464.5	37.1	810	14	AAH34428	Sequence encoded b
37	464.5	37.1	810	18	AAH31169	Plasminogen protei
38	464.5	37.1	810	20	AAH08685	Human plasminogen
39	464.5	37.1	810	20	AAV02114	SEQ ID 77 of W0991
40	464.5	37.1	810	21	AAH82690	Human plasminogen
41	464.5	37.1	810	21	AAH53867	Amino acid sequenc
42	464.5	37.1	810	23	ABH81496	Human plasminogen
43	464.5	37.1	810	23	ABH83795	Human plasminogen.
44	462	36.9	812	16	AAH83959	Complete mouse pla
45	462	36.9	812	17	AAW07585	Murine plasminogen

ALIGNMENTS

RESULT 1

AA11711	ID	AA11711	standard; Protein; 264 AA.
XX	AC	AA11711;	
XX	DT	23-OCT-2000	(first entry)
XX	DE	Mouse serine protease BSSP5 (mBSSP5)	SEQ ID NO:4.
XX	KW	BSSP5; serine protease; human; hBSSP5; mouse; mBSSP5; brain;	
XX	KW	diagnostic marker; antibody; transgenic animal; Alzheimer's disease;	
XX	KW	epilepsy; cancer; inflammation; infertility; pancreatitis;	
XX	KW	prostatic hypertrophy.	
OS	Mus	sp.	
XX	PN	WO200031243-A1.	
XX	PD	02-JUN-2000.	
XX	PF	19-NOV-1999;	99WO-JP06473.
XX	PR	20-NOV-1998;	98JP-0347806.
XX	PA	(FUSO) FUSO PHARM IND LTD.	
XX	PI	Uemura H, Okui A, Kominami K, Yamaguchi N, Mitsui S;	
XX	DR	WPI; 2000-400058/34.	
XX	DR	N-PSDB; AAA61734.	
XX	PT	Serine proteases BSSP5, useful in detecting homologs, mutants and	

11	481.5	38.5	242	7	AAP60061	Sequence of human
12	471.5	37.7	252	7	AAP60058	Sequence of human
13	471	37.6	279	22	ABG20513	Novel human diago
14	469	37.5	269	7	AAP61723	Human elastase II.
15	455.5	37.2	791	18	AAW34285	Human plasminogen.
16	465.5	37.2	810	23	ABB83470	Human plasminogen.
17	465.5	37.2	810	23	ABB75939	Human plasminogen.
18	465.5	37.2	1220	22	AAU32858	Novel human secret
19	464.5	37.1	230	20	AAV31159	Human plasminogen
20	464.5	37.1	249	23	ABB81497	Human microplasmin
21	464.5	37.1	348	23	ABB81498	Human miniplasmin
22	464.5	37.1	437	19	AAW51457	Human plasminogen
23	464.5	37.1	546	14	AAAR34427	Sequence of tissue
24	464.5	37.1	790	15	AAR60519	Human 'Glu' plasmi
25	464.5	37.1	790	22	AAB36562	Mammalian kringle
26	464.5	37.1	791	21	AAB01887	Human plasminogen,
27	464.5	37.1	791	21	AAV95889	Human plasminogen,
28	464.5	37.1	791	21	AAV50867	Human plasminogen
29	464.5	37.1	791	22	AAE67223	Anino acid sequenc
30	464.5	37.1	791	23	AAE23660	Human plasminogen
31	464.5	37.1	810	11	AAR08065	Human plasminogen
32	464.5	37.1	810	12	AAAR13219	R561G human plasmi
33	464.5	37.1	810	12	AAAR12406	R561E human plasmi
34	464.5	37.1	810	12	AAAR13220	R561S human plasmi
35	464.5	37.1	810	12	AAAR13221	Human plasminogen
36	464.5	37.1	810	14	AAAR34428	Sequence encoded b
37	464.5	37.1	810	18	AAW31169	Plasminogen protei
38	464.5	37.1	810	20	AAV08685	Human plasminogen
39	464.5	37.1	810	20	AAV02114	SEQ ID 77 of WO991
40	464.5	37.1	810	21	AAV82690	Human plasminogen
41	464.5	37.1	810	21	AAV53867	Anino acid sequenc
42	464.5	37.1	810	21	AAV53867	Human plasminogen
43	464.5	37.1	810	23	ABB81496	Human plasminogen.
44	462	36.9	812	16	AAR83959	Complete mouse pla
45	462	36.9	812	17	AAW07585	Murine plasminogen

ALIGNMENTS

RESULT 1

AAAB11711

ID AAAB11711 standard; Protein: 264 AA.

XX AAAB11711;

AC AAAB11711;

XX DT 23-OCT-2000 (first entry)

XX Mouse serine protease BSSP5 (mBSSP5) SEQ ID NO:4.

DE BSSP5: serine protease; human; hbSSP5; mouse; mBSSP5; brain;

KW diagnostic marker; antibody; transgenic animal; Alzheimer's disease;

KW epilepsy; cancer; inflammation; infertility; pancreatitis;

KW prostatic hypertrophy.

XX OS Mus sp.

XX WO200031243-A1.

XX PD 02-JUN-2000.

XX PF 19-NOV-1999; 99WO-JP06473.

XX PR 20-NOV-1998; 98JP-0347806.

XX PA (FUSO) FUSO PHARM IND LTD.

XX PI Uemura H, Okui A, Kominami K, Yamaguchi N, Mitsui S;

XX WPI; 2000-400058/34.

DR N-PSDB; AAA61734.

XX PT Serine proteases BSSP5, useful in detecting homologs, mutants and

PT polymorphic variants as markers for diagnosis of e.g. Alzheimer's
 PT disease, epilepsy, cancer and inflammation, using blood, urine,
 PT pancreas or other tissues
 XX
 XX Claim 3; Page 55-56; 70pp; Japanese.
 PS
 CC The invention relates to novel serine proteases designated BSSP5
 CC (AAB11710-B11711), and to nucleic acids encoding them (AAA61733-A61734).
 CC The invention also relates to vectors and transformants comprising BSSP5
 CC nucleic acids; transgenic animals in which the expression level of BSSP5
 CC can be varied; and an mBSSP5 knockout mouse. The invention additionally
 CC encompasses anti-BSSP5 antibodies and methods of production of such
 CC antibodies, methods of BSSP5 detection using the antibodies, and the
 CC use of BSSP5 proteins or fragments as diagnostic markers for certain
 CC medical conditions, e.g., pancreatitis. A method for detecting
 CC pancreatitis comprising measuring BSSP5 concentration in the blood or
 CC urine, and a pancreatitis diagnostic agent containing an anti-BSSP5
 CC antibody is also disclosed. Nucleotides encoding BSSP5 were initially
 CC isolated in a human brain cDNA library using degenerate PCR primers
 CC (AAA61744-A61745) based on conserved regions of serine proteases. The
 CC BSSP5 serine proteases and nucleotides encoding them are useful in
 CC detecting homologues, mutants and polymorphic variants in biological
 CC samples (e.g., blood, urine, brain, prostate gland, placenta, testis,
 CC pancreas and spleen) as diagnostic markers for conditions such as
 CC Alzheimer's disease, epilepsy, cancer, inflammation, infertility,
 CC pancreatitis and prostatic hypertrophy. Sequence AAB11710 represents
 CC human BSSP5 (hBSSP5), and sequence AAB11711 represents murine BSSP5
 CC (mBSSP5).
 XX
 XX Sequence 264 AA;
 SQ
 Query Match 100.0%; Score 1252; DB 21; Length 264;
 Best Local Similarity 100.0%; Pred. No. 3e-102;
 Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IVNGENAVPGSWPWQVSLQDNTGFHFCGGLISPNVWVTAACHQVTPGRHFVVLGEYDRS 60
 DB 34 IVNGENAVPGSWPWQVSLQDNTGFHFCGGLISPNVWVTAACHQVTPGRHFVVLGEYDRS 93
 QY 61 SNAEPQVLSIARAIHPNNANTMNDLTLKLSAPRYTAQVSPVCLASTNEALPSGL 120
 DB 94 SNAEPQVLSIARAIHPNNANTMNDLTLKLSAPRYTAQVSPVCLASTNEALPSGL 153
 QY 121 TCVTGWRISGVGNVTPARLQVPLVTVNOCROYWGARITDAMICAGGSGASSCQGD 180
 DB 154 TCVTGWRISGVGNVTPARLQVPLVTVNOCROYWGARITDAMICAGGSGASSCQGD 213
 QY 181 SGGPLVCQKGNWVLIGIVSWGKNCNIQAPAMYTRVSKFSTWVNOVMAYN 231
 DB 214 SGGPLVCQKGNWVLIGIVSWGKNCNIQAPAMYTRVSKFSTWVNOVMAYN 264
 RESULT 2
 AAB11710
 ID AAB11710 standard; Protein; 264 AA.
 XX
 AC AAB11710;
 AC
 DT 23-OCT-2000 (first entry)
 XX
 DE Human serine protease BSSP5 (hBSSP5) SEQ ID NO:2.
 KW BSSP5; serine protease; human; hBSSP5; mouse; mBSSP5; brain;
 KW diagnostic marker; antibody; transgenic animal; Alzheimer's disease;
 KW epilepsy; cancer; inflammation; infertility; pancreatitis;
 KW prostatic hypertrophy.
 XX
 OS Homo sapiens.
 XX
 PN WO200031243-A1.
 XX
 PD 02-JUN-2000.
 XX

PF 19-NOV-1999; 99WO-JP06473.
 XX
 PR 20-NOV-1998; 98JP-0347806.
 XX
 PA (FUSO) FUSO PHARM IND LTD.
 XX
 PI Uemura H, Okui A, Kominami K, Yamaguchi N, Mitsui S;
 XX
 DR WPI: 2000-400058/34.
 DR N-PSDB; AAA61733.
 XX
 PT Serine proteases BSSP5, useful in detecting homologs, mutants and
 PT polymorphic variants as markers for diagnosis of e.g. Alzheimer's
 PT disease, epilepsy, cancer and inflammation, using blood, urine,
 PT pancreas or other tissues
 XX
 PS Claim 1; Page 51-52; 70pp; Japanese.
 CC The invention relates to novel serine proteases designated BSSP5
 CC (AAB11710-B11711), and to nucleic acids encoding them (AAA61733-A61734).
 CC The invention also relates to vectors and transformants comprising BSSP5
 CC nucleic acids; transgenic animals in which the expression level of BSSP5
 CC can be varied; and an mBSSP5 knockout mouse. The invention additionally
 CC encompasses anti-BSSP5 antibodies and methods of production of such
 CC antibodies, methods of BSSP5 detection using the antibodies, and the
 CC use of BSSP5 proteins or fragments as diagnostic markers for certain
 CC medical conditions, e.g., pancreatitis. A method for detecting
 CC pancreatitis comprising measuring BSSP5 concentration in the blood or
 CC urine, and a pancreatitis diagnostic agent containing an anti-BSSP5
 CC antibody is also disclosed. Nucleotides encoding BSSP5 were initially
 CC isolated in a human brain cDNA library using degenerate PCR primers
 CC (AAA61744-A61745) based on conserved regions of serine proteases. The
 CC BSSP5 serine proteases and nucleotides encoding them are useful in
 CC detecting homologues, mutants and polymorphic variants in biological
 CC samples (e.g., blood, urine, brain, prostate gland, placenta, testis,
 CC pancreas and spleen) as diagnostic markers for conditions such as
 CC Alzheimer's disease, epilepsy, cancer, inflammation, infertility,
 CC pancreatitis and prostatic hypertrophy. Sequence AAB11710 represents
 CC human BSSP5 (hBSSP5), and sequence AAB11711 represents murine BSSP5
 CC (mBSSP5).
 XX
 XX Sequence 264 AA;
 SQ
 Query Match 88.8%; Score 1112; DB 21; Length 264;
 Best Local Similarity 85.7%; Pred. No. 6.3e-90;
 Matches 198; Conservative 22; Mismatches 11; Indels 0; Gaps 0;
 QY 1 IVNGENAVPGSWPWQVSLQDNTGFHFCGGLISPNVWVTAACHQVTPGRHFVVLGEYDRS 60
 DB 34 IVNGENAVPGSWPWQVSLQDNTGFHFCGGLISQSWVWVTAACHCNVSPGRHFVVLGEYDRS 93
 QY 61 SNAEPQVLSIARAIHPNNANTMNDLTLKLSAPRYTAQVSPVCLASTNEALPSGL 120
 DB 94 SNAEPQVLSIARAIHPNNANTMNDLTLKLSAPRYTAQVSPVCLASTNEALPSGL 153
 QY 121 TCVTGWRISGVGNVTPARLQVPLVTVNOCROYWGARITDAMICAGGSGASSCQGD 180
 DB 154 TCVTGWRISGVGNVTPARLQVPLVTVNOCROYWGARITDAMICAGGSGASSCQGD 213
 QY 181 SGGPLVCQKGNWVLIGIVSWGKNCNIQAPAMYTRVSKFSTWVNOVMAYN 231
 DB 214 SGGPLVCQKGNWVLIGIVSWGKNCNVRAPAVYTRVSKFSTWVNOVMAYN 264
 RESULT 3
 AAB54077
 ID AAB54077 standard; Protein; 192 AA.
 XX
 AC AAB54077;
 XX
 DT 09-MAR-2001 (first entry)
 XX
 DE Human pancreatic cancer antigen protein sequence SEQ ID NO:529.

XX Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
KW detection; diagnosis; identification; cytostatic; neuroprotective;
KW nontropic; immunomodulatory; relaxant; contraceptive; gynaecological;
KW antinflammatory; cardiant; gene therapy; chromosome mapping;
KW linkage analysis; tissue identification; tissue typing; forensic;
KW neural; immune system; muscular; reproductive; gastrointestinal;
KW pulmonary; cardiovascular; renal; proliferative.
XX
OS Homo sapiens.
XX
XX WO200055320-A1.
PN
XX 21-SEP-2000.
PD
XX 08-MAR-2000; 2000WO-US05989.
PF
XX 12-MAR-1999; 99US-0124270.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Rosen CA, Ruben SM;
PI
XX
XX WPI; 2000-579444/54.
DR N-PSDB; AAC98842.
XX
XX New nucleic acid that is a pancreatic cancer antigen for preventing,
PT treating, or ameliorating a medical condition, particular pancreatic
PT cancer, or for use in assays for diagnosing a pathological condition -
PT
XX Claim 11; Page 966; 1379pp; English.
PS
XX AAC98773 to AAC9231 encode the human pancreatic cancer associated
CC proteins, called pancreatic cancer antigens, given in AAB54008 to
CC AAB54466. The human pancreatic cancer antigens have cytostatic,
CC neuroprotective, nontropic, immunomodulatory, relaxant, contraceptive,
CC gynaecological, cardiant and antiinflammatory activities, and can be used
CC in gene therapy. The polynucleotide and proteins can be used for
CC preventing, treating, or ameliorating a medical condition or in assays
CC for diagnosing a pathological condition or a susceptibility to one in a
CC subject. Binding partners to the proteins and the activity of the
CC proteins can be identified. The pancreatic cancer antigens can be used to
CC detect, treat or prevent pancreatic disorders, especially cancer.
CC Agonists and antagonists to the antigens can be screened for. The
CC pancreatic cancer antigen polynucleotides can be used to design nucleic
CC acid hybridisation probes that can be used in chromosome mapping, linkage
CC analysis, tissue identification and/or typing and a variety of forensic
CC and diagnostic methods. The proteins can be used to generate antibodies
CC which are used to purify, detect and target the polypeptides, including
CC both in vivo and in vitro diagnostic and therapeutic methods. The
CC proteins can be used to treat or prevent neural, immune system, renal or
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC proliferative disorders. AAC9232 to AAC9240 and AAB54467 represent
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 192 AA;
Query Match 60.4%; Score 756; DB 21; Length 192;
Best Local Similarity 84.4%; Pred. No. 9.2e-59;
Matches 135; Conservative 16; Mismatches 9; Indels 0; Gaps 0;
QY 1 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAACQVTPGRHFVVLGEYDRS 60
Db 30 IVNGENAVLGSWPQVSLQDSSGFHFCGSLISQSNWVYTAACNVSPGRHFVVLGEYDRS 89
QY 61 SNAEPQVLSIARATHPHNNANTMNDTLKLPASARYTAQVSPVCLASTNEALPSGL 120
Db 90 SNAEPQLVLSVRAITHPSWNTNDVTLKLPASPAQYTTTRISPVCLASNEALTEGL 149
QY 121 TCVTGWRISGVNVTPLAQVVLPLVTVNCQRYMGA 160
Db 150 TCVTGWRISGVNVTPLAQVVLPLVTVNCQRYMGS 189

RESULT 4

AAU82738
ID AAU82738 standard; Protein; 263 AA.
XX
AC AAU82738;
XX
DT 23-APR-2002 (first entry)
XX
DE Amino acid sequence of novel human protease #37.
XX
KW Human; protease; cancer; immune-related disorder; cardiovascular disease;
KW neuronal-associated disease; metabolic disorder; inflammatory disorder;
KW nervous system disorder; sexual dysfunction; pain; mood disorder;
KW hypertension; psychotic disorder; neurological disorder; dyskinesia;
KW viral infection; human immunodeficiency virus; HIV; non-viral infection;
KW ocular disease; cytostatic; enzyme.
XX
OS Homo sapiens.
XX
XX WO200200860-A2.
PN
XX 03-JAN-2002.
PD
XX 26-JUN-2001; 2001WO-US20171.
PF
XX 26-JUN-2000; 2000US-214047P.
PR
XX (SUGE-) SUGEN INC.
PA
XX
XX Plowman G, Whyte D, Sudarsanam S, Manning G, Caenepeel S;
PI Charyczak G;
XX
XX WPI; 2002-139913/18.
DR N-PSDB; ABR31780.
XX
XX Nucleic acids encoding novel human proteases, useful for useful for
PT treating diseases and disorders such as cancers, immune-related
PT diseases and disorders, cardiovascular disease (e.g. restenosis) and
PT inflammatory disorders -
XX
XX Claim 6; Fig 2M; 313pp; English.
XX
XX The present invention relates to the isolation of novel human
CC proteases, and the nucleic acids encoding them. The sequences of
CC the invention are useful for treating diseases and disorders such as
CC cancers (e.g. breast, colon, lung), immune-related diseases and disorders
CC (e.g. inflammatory diseases and asthma), cardiovascular diseases
CC (e.g. restenosis and coronary thrombosis), brain or neuronal-associated
CC diseases, metabolic disorders (e.g. diabetes, obesity), inflammatory
CC disorders (e.g. rheumatoid arthritis and psoriasis), central or
CC peripheral nervous system diseases, migraines, pain, sexual dysfunction,
CC mood disorders, attention disorders, cognitive disorders, hypotension,
CC hypertension, psychotic disorders, neurological disorders
CC (e.g. Alzheimer's disease, Parkinson's disease) and dyskinesias.
CC The nucleic acids and polypeptides are also useful for treating viral
CC infections caused by human immunodeficiency virus (HIV), and non-viral
CC infections such as ocular disease (e.g. glaucoma) and macular
CC degeneration. AAU82702-AAU82760 represent the novel human proteases of
CC the invention.
XX
SQ Sequence 263 AA;
Query Match 60.3%; Score 755.5; DB 23; Length 263;
Best Local Similarity 58.4%; Pred. No. 1.5e-58;
Matches 135; Conservative 35; Mismatches 60; Indels 1; Gaps 1;
QY 1 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAACQVTPGRHFVVLGEYDRS 60
Db 34 IVNGEDAVPGSWPQVSLQDKTGFHFCGSLISEDWVYTAACGVRTS-DVVVAGEFDG 92
QY 61 SNAEPQVLSIARATHPHNNANTMNDTLKLPASARYTAQVSPVCLASTNEALPSGL 120
Db 150 SNAEPQVLSIARATHPHNNANTMNDTLKLPASARYTAQVSPVCLASTNEALPSGL 120

Db 93 SDEENIQVLKIAKFKNPKFSILTVNNDITLLKLTATPAREFSQTVSAVCLPSADDDFFAGT 152
QY 121 TCVTTGWRISGVGNVTPARLQOVVPLVTVNOCROYWGARITDAMICAGSGASSCGGD 180
Db 153 LCATTGKGKYNANKTPDKLQQAALPLLSNAECKSKWGRITDVMICAGSGVSSCMGD 212
QY 181 SGGPLVCQKGNWTLIGIVSGTKNCNIQAPAMTRYSKFSTWINOVMAYN 231
Db 213 SGGPLVCQKGAWTLVGIVSGWSTCSTTPPAVYARVTKLIPWQKILAA 263

RESULT 5
AAB98504
ID AAB98504 standard; Protein; 231 AA.
XX AC AAB98504;
DT 03-AUG-2001 (first entry)
DE Human chymotrypsin serine protease catalytic domain.
KW Human; TADG-15; cytostatic; vaccine; ovarian tumour; cancer;
KW tumour antigen-derived gene 15; serine protease; chymotrypsin.
OS Homo sapiens.
XX WO200129056-A1.
XX 26-APR-2001.
XX 20-OCT-2000; 2000WO-US29095.
XX 20-OCT-1999; 99US-0421213.
XX (UYAR-) UNIV ARKANSAS.
XX O'Brien TJ, Tanimoto H;
XX WPI; 2001-381031/40.
XX Novel extracellular serine protease, termed tumor antigen-derived gene
PT 15 protein overexpressed in carcinomas and DNA encoding it, for
PT diagnosis, treatment, prevention of cancer, particularly breast,
PT ovarian cancer
XX Example 10; Fig 1; 130pp; English.
XX The present invention relates to human tumour antigen-derived gene 15
CC (TADG-15) protein and coding sequence (see AAB23601 and AAB98500).
CC TADG-15 is an extracellular serine protease. It was found that TADG-15 is
CC over-expressed in ovarian tumours. TADG-15 protein or its fragments of
CC 9-20 residues that lack TADG-15 protease activity are useful for
CC vaccinating an individual against TADG-15, having, suspected of having or
CC at risk of getting cancer. Furthermore, the TADG-15 gene can be used as a
CC diagnostic or therapeutic target in cancer. The present sequence was used
CC in a sequence homology alignment with the catalytic domain of TADG-15.
XX
SQ Sequence 231 AA;

Query Match 59.9%; Score 750.5; DB 22; Length 231;
Best Local Similarity 58.0%; Pred. No. 3.5e-58;
Matches 134; Conservative 35; Mismatches 61; Indels 1; Gaps 1;

QY 1 IVNGEAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAACHQVTPGRHFVVLGEYDRS 60
Db 2 IVNGEDAVPGSWPQVSLQDKTGFHFCGSLISDNWVYTAACHGVRTS-DVVVAGEFDQG 60
QY 61 SNAEPQVLSIARAITHPNNANMNDLTLLKLASPARYTAQVSPVCLASTNEALPSGL 120
Db 1 SDEENIQVLKIAKFKNPKFSILTVNNDITLLKLTATPAREFSQTVSAVCLPSADDDFFAGT 120
QY 121 TCVTTGWRISGVGNVTPARLQOVVPLVTVNOCROYWGARITDAMICAGSGASSCGGD 180

Db 121 LCATTGKGKYNANKTPDKLQQAALPLLSNAECKSKWGRITDVMICAGSGVSSCMGD 180
QY 181 SGGPLVCQKGNWTLIGIVSGTKNCNIQAPAMTRYSKFSTWINOVMAYN 231
Db 181 SGGPLVCQKGAWTLVGIVSGWSTCSTSPGVYARVTKLIPWQKILAA 231

RESULT 6
AAY99596
ID AAY99596 standard; protein; 245 AA.
XX AC AAY99596;
DT 13-SEP-2000 (first entry)
DE Bovine chymotrypsinogen A.
KW Bovine; plasminogen activator; cardiant; thrombolytic;
KW heart attack; stroke; blood clotting disorder.
OS Bos taurus.
XX WO200032759-A1.
XX 08-JUN-2000.
XX 06-MAY-1999; 99WO-US09991.
XX 02-DEC-1998; 98US-0110588.
XX (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX Lin X, Zhang XC, Tang JJN;
XX WPI; 2000-422975/36.
XX Polypeptide with plasminogen activator activity useful as thrombolytic
PT agent for treating blood clot disorders e.g. heart attack, comprises 10
PT amino acid peptide fragment for recognition or activation of
PT plasminogen
XX Disclosure; Page 35-36; 41pp; English.
XX The present sequence is bovine chymotrypsinogen. It was included in
CC a review of sequence homologies of several plasminogen activators.
CC Plasminogen is the principal serine protease zymogen in the
CC extracellular fluids of vertebrates. Its active form, plasmin, is
CC implicated in pericellular proteolysis associated with a wide range of
CC physiological and pathological processes. Plasminogen expression is
CC regulated by plasminogen activators which hydrolyse a peptide bond in
CC plasminogen to convert it to plasmin or form tight binding complexes
CC with plasminogen to spontaneously convert it to plasmin. The sequence
CC homology analysis has identified a six amino acid peptide involved in
CC plasminogen activation. This peptide is particularly useful when inserted
CC between amino acid residues 644 and 645 of full length human
CC plasminogen. Novel plasminogen activators have been made based upon the
CC plasminogen activation/recognition site of plasminogen binding proteins.
CC The polypeptides are useful in preparing thrombolytic agents for treating
CC blood clotting disorders such as heart attack.
XX
SQ Sequence 245 AA;

Query Match 58.1%; Score 727.5; DB 21; Length 245;
Best Local Similarity 56.7%; Pred. No. 4e-56;
Matches 131; Conservative 32; Mismatches 67; Indels 1; Gaps 1;

QY 1 IVNGEAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAACHQVTPGRHFVVLGEYDRS 60
Db 16 IVNGEDAVPGSWPQVSLQDKTGFHFCGSLISDNWVYTAACHGVRTS-DVVVAGEFDQG 74
QY 61 SNAEPQVLSIARAITHPNNANMNDLTLLKLASPARYTAQVSPVCLASTNEALPSGL 120
Db 75 SSEEKIQKIAKFKNPKSKYNSLTINNDITLLKLTAAFSQTVSAVCLPSASDDFAAGT 134

Qy 230 YN 231
—
Db 268 NN 269

RESULT 9
AAR29621
ID AAR29621 standard; Protein; 253 AA.

```

XX AAR29621;
XX 30-APR-1993 (first entry)
XX Porcine pancreatic elastase III.
XX Pig; swine; bile acid; gall bladder; bile acid secretion promoters;
XX liver function improvers; ss.
XX Sus scrofa domestica.
XX JP04325090-A.
XX 13-NOV-1992.
XX 23-APR-1991; 91JP-0092069.
XX 23-APR-1991; 91JP-0092069.
XX (SANY ) SANKYO CO LTD.
XX WPI; 1992-428826/52.
XX N-PSDB; AAQ31724.
XX Pig pancreas elastase protein - used in bile acid secretion
XX promoters and liver function improvers
XX Claim 1; Page 8; 10pp; Japanese.
XX This sequence represents porcine pancreatic elastase III. It was
XX expressed in E. coli V421 using the expression vector pELE001.
XX It may be used in bile acid secretion promoters and liver function
XX improvers.
XX Sequence 253 AA;
Query Match 40.0%; Score 500.5; DB 13; Length 253;
Best Local Similarity 43.2%; Pred. No. 3.9e-36;
Matches 104; Conservative 42; Mismatches 82; Indels 13; Gaps 9;
QY 1 IVNGENAVPGSPWQVSLQ--DNTGF--HFCGGSLISPNWVYTAHQCQVTPGRHFVVLGEY 57
DB 12 VVNGEDAVPYSPWQVSLQEKNGVFQHTCGGSLIAPDWLTAGHCISSSLTYQVVLGEY 71
QY 58 DRSSNAEPQVLSI--ARAITHPNWNTMN--NDLTLLKLSPARYTAQVSPVCLASTN 113
DB 72 DRSENEGFEQVIPINAGDLFVHPRWNSVCVSCGNDIALVKLSAQIGDKVQLACLPPAG 131
QY 114 EALPSGLTCTVTTGWRISGVGNVTPARLQOVPLVTVNOCRO--YWGARITDAMICAGG 171
DB 132 DILPNDTPCIVSGRLTYNGPL--PKLQALLPVVDYQHCKWDWNGSTVKTMTWCAGG 190
QY 172 SGASSCGDGGPLVCQKGN--TWVLGIVSW--GTKNCN--IQAPAMYTRVSKFSTWINQVM 228
DB 191 DIRSGCNGDGGPLNCPAADSGMQVHGVSFVSAYGNTLKKPTVTRTSAFIDWIEEII 250
QY 229 A 229
DB 251 A 251
RESULT 10
AAP60062
ID AAP60062 standard; Protein; 269 AA.
XX
XX AAP60062;
XX 06-SEP-1991 (first entry)
XX Sequence of human pancreatic elastase IIA encoded on pH2E2.
XX Enzyme; serum lipoprotein metabolism.

```

```

XX OS Homo sapiens.
XX EP198645-A.
XX 22-OCT-1986.
XX 07-APR-1986; 86EP-0302557.
XX 23-OCT-1985; 85JP-0236686.
XX 05-APR-1985; 85JP-0072308.
XX 27-APR-1985; 85JP-0091986.
XX 26-JUL-1985; 85JP-0163964.
XX 02-DEC-1985; 85JP-0271128.
XX (SANY ) SANKYO KK.
XX Takiguchi Y, Tani T, Kawashima I, Erukawa H, Ohmine T;
XX Ohsumi J;
XX WPI; 1986-280300/43.
XX N-PSDB; AAN60076.
XX New genetically engineered human pancreatic elastase - obt'd.
XX using hosts modified DNA coding for enzyme
XX Example; Page 14-15; 45pp; English.
XX The genetically engineered prod. can eliminate the dependency on
XX human pancreas supplies for the elastase, and avoids antibody
XX formation and possibility of anaphylaxis using porcine elastase.
XX Sequence 269 AA;
Query Match 38.8%; Score 486; DB 7; Length 269;
Best Local Similarity 43.0%; Pred. No. 8e-35;
Matches 104; Conservative 41; Mismatches 85; Indels 12; Gaps 8;
QY 1 IVNGENAVPGSPWQVSLQDNTG--FHCGGSLISPNWVYTAHQCQVTPGRHFVVLGEY 57
DB 29 VVGEEARPNWQVSLQYSSNGKWTHTCGGSLIANSWVLTAAHCISSTRTVVLGRH 88
QY 58 DRSSNAEPQVLSIARAITHPNWNTMN--NDLTLLKLSPARYTAQVSPVCLASTNEA 115
DB 89 NLVVAESGSLAVSVSKIVVHKDWNQISKGNIDIALKLANPVSLTDKIQLACLPAGTI 148
QY 116 LPSGLTCTVTTGWRISGVGNVTPARLQOVPLVTVNOCRO--YWGARITDAMICAGGSG 173
DB 149 LPNYPCTVTTGWRISGVGNVTPARLQOVPLVTVNOCRO--YWGARITDAMICAGGSG 207
QY 174 A--SSCGDGGPLVCQKGN--TWVLGIVSWGK--NCN--IQAPAMYTRVSKFSTWINQVM 229
DB 208 VISCSGNGDGGPLNCPAADSGMQVHGVSFVSAYGNTLKKPTVTRTSAFIDWIEEII 267
QY 230 YN 231
DB 268 NN 269
RESULT 11
AAP60061
ID AAP60061 standard; Protein; 242 AA.
XX
XX AAP60061;
XX 06-SEP-1991 (first entry)
XX Sequence of human pancreatic elastase IIIB.
XX Enzyme; serum lipoprotein metabolism.
XX Homo sapiens.
XX

```

```
PN EPI98645-A.
XX
PD 22-OCT-1986.
XX
PF 07-APR-1986; 86EP-0302557.
XX
PR 23-OCT-1985; 85JP-0236686.
XX
PR 05-APR-1985; 85JP-0072308.
PR 27-APR-1985; 85JP-0091986.
PR 26-JUL-1985; 85JP-0163964.
PR 02-DEC-1985; 85JP-0271128.
XX
PA (SANY ) SANKYO KK.
XX
XX Takiguchi Y, Tani T, Kawashima I, Eurukawa H, Ohmine T;
PI Ohsumi J;
PI
XX
XX WPI; 1986-280300/43.
DR N-PSDB; AAN60075.
DR
XX
XX New genetically engineered human pancreatic elastase - obt'd.
PT using hosts modified DNA coding for enzyme
XX
XX Claim 2; Page 8; 45pp; English.
XX
XX The genetically engineered prod. can eliminate the dependency on
CC human pancreas supplies for the elastase, and avoids antibody
CC formation and possibility of anaphylaxis using porcine elastase.
XX
XX
XX Sequence 242 AA;
SQ
Query Match 38.5%; Score 481.5; DB 7; Length 242;
Best Local Similarity 42.7%; Pred. No. 1.7e-34;
Matches 103; Conservative 39; Mismatches 86; Indels 13; Gaps 9;
Qy 1 IVNGENAVPGSWPQVLSIQ-DNTG--FHFCGSLISPNVYTAHCQVTPGRHFVVLGEY 57
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1 VVNGEDAVPYSPWPQVLSIQEYKSGSFYHFCGSLIAPDMVYTAGHCISSSRTYQVVLGEY 60
Qy 58 DRSSNAEPVQVLSI--ARAITHPNNTNM--NNDLTLLKLSPARYTAQVSPVCLASTN 113
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
61 DRVKEGPEQVPIPSNGDLFVHPLNRSVACGNDIALIKLSRQGLDGVQLASLPAG 120
Qy 114 EALPSGLTCVTGMRISGVGNVTPARLQVVLPLVTYNOCRQ--YWGARITDAMICAGG 171
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
121 DILNETECYITGWRGLVTNGPL-PDKLQEALLPVVDYEHCSRNNWGWSSVKKTWVCAGG 179
Qy 172 SGASSCQDGGPLVC-OKGNTWVLIGIVSW-GTKNCNI-QAPAMYTRVSKFSTWINQVM 228
Db | ||||| | : | : | : | : | : | : | : | : | : | : | : | : | :
180 DIRSCNGDGGPLNCPTEDGGWQVHGVTSEVSAFGCNRTRKPTVTRVSADFIDWIEETI 239
Qy 229 A 229
Db 240 A 240
RESULT 12
AAP60058
ID AAP60058 standard; Protein; 252 AA.
XX
AC AAP60058;
XX
XX 06-SEP-1991 (first entry)
DT
DE Sequence of human pancreatic elastase IIA.
XX
XX Enzyme; serum lipoprotein metabolism.
XX
XX Homo sapiens.
XX
XX EPI98645-A.
XX
XX 22-OCT-1986.
PD
```

```
XX
PF 07-APR-1986; 86EP-0302557.
XX
XX 23-OCT-1985; 85JP-0236686.
PR 05-APR-1985; 85JP-0072308.
PR 27-APR-1985; 85JP-0091986.
PR 26-JUL-1985; 85JP-0163964.
PR 02-DEC-1985; 85JP-0271128.
XX
PA (SANY ) SANKYO KK.
XX
XX Takiguchi Y, Tani T, Kawashima I, Eurukawa H, Ohmine T;
PI Ohsumi J;
PI
XX
XX WPI; 1986-280300/43.
DR N-PSDB; AAN60072.
DR
XX
XX New genetically engineered human pancreatic elastase - obt'd.
PT using hosts modified DNA coding for enzyme
XX
XX Claim 2; Page 3-4; 45pp; English.
XX
XX The genetically engineered prod. can eliminate the dependency on
CC human pancreas supplies for the elastase, and avoids antibody
CC formation and possibility of anaphylaxis using porcine elastase.
XX
XX
XX Sequence 252 AA;
SQ
Query Match 37.7%; Score 471.5; DB 7; Length 252;
Best Local Similarity 42.6%; Pred. No. 1.4e-33;
Matches 103; Conservative 41; Mismatches 85; Indels 13; Gaps 9;
Qy 1 IVNGENAVPGSWPQVLSIQDNTG---FHFCGSLISPNVYTAHCQVTPGRHFVVLGEY 57
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
13 VVGGEARPNWPQVLSIQYSSNGKWIYTCGSLIANSWVLTAAHCISSTRYVGLGRH 72
Qy 58 DRSSNAEPVQVLSIARAITHPNNTNM--NNDLTLLKLSPARYTAQVSPVCLASTNEA 115
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
73 NLYVAESGLAVSVSKIVVHKDWSNQISKGNIDALLKLANPVSLTDKIQLACLPPAGTI 132
Qy 116 LPSGLTCVTGMRISGVGNVTPARLQVVLPLVTYNOCRQ--YWGARITDAMICAGSG 173
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
133 LPNYPVCYVTGWRGLVTNGAV-PDVLOQRLVAVAYATCSSAWGSSVYKTS-ICAGG 190
Qy 174 A-SSCQDGGPLVCQKGN-TWVLIGIVSWGTFK-NCN-IOAPAMYTRVSKFSTWINQVMA 229
Db | ||||| | : | : | : | : | : | : | : | : | : | : | : | : | :
191 VISSCNGDGGPLNCQASDGRQVHGIVSFGSLGNCNYHKKPSVTRVSNYIDWINSVIA 250
Qy 230 YN 231
Db 251 NN 252
RESULT 13
ABG20513
ID ABG20513 standard; Protein; 279 AA.
XX
AC ABG20513;
XX
XX 13-FEB-2002 (first entry)
DT
DE Novel human diagnostic protein #20504.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
OS
XX WO200175067-A2.
PN
XX 11-OCT-2001.
PD
XX 30-MAR-2001; 2001WO-US08631.
PF
```

XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 PA Drmanac RT, Liu C, Tang YT;
 XX WPI: 2001-639362/73.
 PI N-PSDB; AAS84700.
 DR New isolated polynucleotide and encoded polypeptides, useful in
 DR diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PT Claim 20; SEQ ID No 50872; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABC0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 279 AA;
 SQ

Query Match 37.6%; Score 471; DB 22; Length 279;
 Best Local Similarity 41.0%; Pred. No. 1.7e-33;
 Matches 102; Conservative 39; Mismatches 84; Indels 24; Gaps 9;
 QY 1 IVNGENAVPGSWPQVSLQDNTG---FHFCGGSLISPNWVTAHC-----QVTPGRH 50
 Db 37 VVGGEARPNNSWPQVSLQYSSNGKWHYTCGGSLIANSWVLTAAHCISSTRYRVLGRH 96
 QY 51 FVYLGEYDRSSNAEPVQVLSIARAITHPNWNTM--NDLTLLKLASPARYTAQVSPVC 108
 Db 97 NLYVAE-----SGSLAMSSVSKIVVHKDNFQISKGNIDALLKLANPVSULTDKIQLAC 150
 QY 109 LASTNEALPSGLTCVTTGWRISGVGNVTPARLQOVVPLVTVNQCRRQ--YWGARITDAM 166
 Db 151 FPPAGTILPNPYPCVXTGLGESCRTNGAVPDVLOQGRLLVVDYATCSSAWNGSSVKTM 210
 QY 167 ICAGGSGA-SSCGDGGPLVCCQKNT--VWLGIVSGWTK-NCN-IOAPAMYTRVSKFST 222
 Db 211 ICAGGSGVSSCGDGGPLNCOASDALMQVHGVISFGSLGNCNYHKPSVFTVRSNYID 270
 QY 223 WINQVMAYN 231
 Db 271 WINSVIAYN 279
 RESULT 14
 AAP61723
 ID AAP61723 standard; Protein; 269 AA.
 XX
 AC AAP61723;
 XX

DT 23-OCT-1991 (first entry)
 XX Human elastase II.
 DE Colibacillus; yeast.
 XX Homo sapiens.
 OS JP61192288-A.
 PN 26-AUG-1986.
 PD 22-FEB-1985; 85JP-0034049.
 PF 22-FEB-1985; 85JP-0034049.
 PR (KIRI) KIRIN BREWERY KK.
 XX WPI: 1986-262894/40.
 DR N-PSDB; AAM60706.
 XX Biologically producing human elastase ii - using DNA chain with
 PT biological prodn. capability for human elastase II.
 PS Claim 1; Fig 1-2; 19pp; Japanese.
 XX The human elastase product may be efficiently expressed from a
 CC transformed host such as colibacillus or yeast, yielding the product
 CC more efficiently than spleen extraction.
 CC See also J61192289.
 XX Sequence 269 AA;
 SQ

Query Match 37.5%; Score 469; DB 7; Length 269;
 Best Local Similarity 42.6%; Pred. No. 2.5e-33;
 Matches 103; Conservative 40; Mismatches 87; Indels 12; Gaps 8;
 QY 1 IVNGENAVPGSWPQVSLQDNTG---FHFCGGSLISPNWVTAHCQVTPGRHVVVLGEY 57
 Db 29 VVGGEARPNNSWPQVSLQYSSNGKWHYTCGGSLIANSWVLTAAHCISSTRYRVLGRH 88
 QY 58 DRSSNAEPVQVLSIARAITHPNWNTM--NDLTLLKLASPARYTAQVSPVCCLASTNEA 115
 Db 89 NLYVAESGSLAVSVKIVVHKDNWNSQISKGNIDALLKLANPVSULTDKIQLACLPAGTI 148
 QY 116 LPSGLTCVTTGWRISGVGNVTPARLQOVVPLVTVNQCRRQ--YWGARITDAMICAGSG 173
 Db 149 LPNNYPCVYTGWRILQTNAGV-PDVLOQGRLLVVDYATCSSAWNGSSVTKMICAGDGG 207
 QY 174 A-SSCGDGGPLVCCQKNTWLI-GIVSWGK-NCN-IOAPAMYTRVSKFSTWINQVMA 229
 Db 208 VISSCNGDGGPLNCOASDGRCEVHGVISFGSLGNCNYHKPSVFTVRSNSIDWINSVIA 267
 QY 230 YN 231
 Db 268 NN 269
 RESULT 15
 AAW34285
 ID AAW34285 standard; protein; 791 AA.
 XX
 AC AAW34285;
 XX
 DT 14-MAY-1998 (first entry)
 XX Human plasminogen.
 DE Plasminogen; human; Kringle 5 peptide; anti-angiogenesis agent; cancer;
 KW metastatic solid tumour; carcinoma; sarcoma; lymphoma; haemangioma;
 KW psoriasis; arthritis; macular degeneration; diabetic retinopathy;
 KW autoimmune disease; ocular disease; capillary proliferation; therapy;
 KW Kringle 5 receptor.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:14:08 ; Search time 4.08696 Seconds
(without alignments)
1663.721 Million cell updates/sec

Title: US-09-856-319B-4_COPY_1_33

Perfect score: 166

Sequence: 1 MLLSLTSLVLLGSSMCGVPAITPALSYNQR 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	166	100.0	264	11 Q9D960	Q9d960 mus musculus
2	166	100.0	264	11 Q9D7P8	Q9d7p8 mus musculus
3	166	100.0	264	11 Q9EQZ8	Q9eqz8 rattus norv
4	166	100.0	264	11 Q9ER05	Q9er05 mus musculus
5	71	42.8	263	11 Q9DC86	Q9dc86 mus musculus
6	71	42.8	263	11 Q9DX88	Q9dx88 mus musculus
7	71	42.8	263	11 Q9CR35	Q9cr35 mus musculus
8	59.5	35.8	261	13 Q9W7Q4	Q9w7q4 paralichthy
9	59	35.5	648	17 Q8THZ0	Q8thz0 methanosarc
10	58	34.9	216	4 Q96R18	Q96r18 homo sapien
11	56	33.7	260	13 Q9W7Q3	Q9w7q3 paralichthy
12	55.5	33.4	140	9 Q8SBN3	Q8sbn3 bacterioph
13	55.5	33.4	441	5 Q96666	Q96666 drosophila
14	55.5	33.4	441	5 Q9V3B9	Q9v3b9 drosophila
15	55.5	33.4	835	17 Q97Z55	Q97z55 sulfolobus
16	55	33.1	242	16 Q8ZLR7	Q8zlr7 salmonella

17	55	33.1	242	16	Q8Z3G0	Q8z3g0 salmonella
18	55	33.1	304	10	Q94AJ4	Q94aj4 arabidopsis
19	55	33.1	465	10	Q9SVB9	Q9sve9 arabidopsis
20	54	32.5	255	10	Q9FW90	Q9fw90 oryza sativ
21	54	32.5	622	16	Q8XT78	Q8xt78 ralstonia s
22	54	32.5	637	2	Q9RPI8	Q9rpi8 desulfitoba
23	53	31.9	671	10	Q9FEA3	Q9fea3 oryza sativ
24	52	31.3	241	16	Q8ZB72	Q8zb72 versinia pe
25	52	31.3	292	10	Q43752	Q43752 citrus sine
26	52	31.3	579	10	Q48542	Q48542 hordeum vul
27	52	31.3	891	12	Q9J052	Q9j052 simian herp
28	51.5	31.0	270	13	Q9I039	Q9i039 gadus morhu
29	51.5	31.0	343	2	Q8RSP6	Q8rsp6 rhizobium l
30	51.5	31.0	663	10	Q04976	Q04976 mangifera 1
31	51	30.7	326	16	Q8ZM02	Q8zmq2 salmonella
32	51	30.7	533	11	Q08569	Q08569 cavia porce
33	51	30.7	757	10	Q8VXA8	Q8vxa8 oryza sativ
34	51	30.7	826	10	Q80739	Q80739 arabidopsis
35	51	30.7	1186	12	Q55767	Q55767 chilo iride
36	50.5	30.4	248	6	Q9N0G1	Q9n0g1 equus cabal
37	50.5	30.4	248	6	Q95L88	Q95l88 equus cabal
38	50.5	30.4	384	16	Q8ZF12	Q8zf12 yersinia pe
39	50.5	30.4	536	11	Q8VHD6	Q8vhd6 mus musculu
40	50.5	30.4	538	4	Q8WUG5	Q8wug5 homo sapien
41	50.5	30.4	1004	16	Q9RJT8	Q9rjt8 streptomyce
42	50	30.1	94	4	Q96F38	Q96f38 homo sapien
43	50	30.1	144	5	Q9V316	Q9v316 drosophila
44	50	30.1	169	4	Q9H7M4	Q9h7m4 homo sapien
45	50	30.1	222	11	Q9JIY6	Q9jiy6 rattus ratt

ALIGNMENTS

RESULT 1

Q9D960 ID Q9D960 PRELIMINARY; PRT; 264 AA.
AC Q9D960;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE 1810004D15RIK protein.
GN CTRL OR 1810004D15RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombert P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.

DR EMBL; AK007333; BAB24967.1; --
DR HSSP; P00766; 4CHA.
DR MEROPS; S01.256; --
DR MGD; MGI:88558; Ctr1.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28151 MW; 1D979719E07C16DE CRC64;

Query Match 100.0%; Score 166; DB 11; Length 264;
Best Local Similarity 100.0%; Pred. No. 2.1e-14;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLLSLTSLVLGSSGCGVPAITPALSYNQR 33
|||||
DB 1 MLLSLTSLVLGSSGCGVPAITPALSYNQR 33
|||||

RESULT 2

ID Q9D7P8 PRELIMINARY; PRT; 264 AA.
AC Q9D7P8;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE 1810004D15RIK protein.
GN CTRL OR 1810004D15RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.

DR EMBL; AK009019; BAB26029.1; --
DR HSSP; P00766; 4CHA.
DR MEROPS; S01.256; --
DR MGD; MGI:88558; Ctr1.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28167 MW; 1D979469A07056C2 CRC64;

Query Match 100.0%; Score 166; DB 11; Length 264;
Best Local Similarity 100.0%; Pred. No. 2.1e-14;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLLSLTSLVLGSSGCGVPAITPALSYNQR 33
|||||
DB 1 MLLSLTSLVLGSSGCGVPAITPALSYNQR 33
|||||

RESULT 3

ID Q9EQZ8 PRELIMINARY; PRT; 264 AA.
AC Q9EQZ8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Chymopasin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=RAT PANCREAS;
RA Sogame Y., Mitsui S., Kataoka K., Kato M., Sakagami J.,
RA Yamaguchi N.;
RT "Molecular cloning of rat chymopasin."
RL Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.

DR EMBL; AB020757; BAB20287.1; --
DR HSSP; P00766; 4CHA.
DR MEROPS; S01.256; --
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28116 MW; F9ED5D210FD3500E CRC64;
Query Match 100.0%; Score 166; DB 11; Length 264;
Best Local Similarity 100.0%; Pred. No. 2.1e-14;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLLSLTSLVLGSSGCGVPAITPALSYNQR 33
|||||
DB 1 MLLSLTSLVLGSSGCGVPAITPALSYNQR 33
|||||

RESULT 4

ID Q9ER05 PRELIMINARY; PRT; 264 AA.
AC Q9ER05;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Chymopasin (Chymotrypsin A CTRA-1).
GN CTRL OR CTRA1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

CC	TRYPSIN FAMILY.					
DR	EMBL; AK003060; BAB22539.1; -.					
DR	HSSP; P00766; 1GCT.					
DR	MEROPS; S01.152; -.					
DR	MGI;1913723; 2200008D09Rik					
DR	InterPro: IPR001314; Chymotrypsin.					
DR	InterPro: IPR001254; Ser_protease_Try.					
DR	Pfam; PF00089; trypsin; 1.					
DR	PRINTS; PR00722; CHYMOTRYPSIN.					
DR	SMART; SM00020; TRYSP_SPC; 1.					
DR	PROSITE; PS50240; TRYPsin_DOM; 1.					
DR	PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.					
DR	PROSITE; PS00135; TRYPsin_SER; 1.					
KW	Hydrolase; Serine protease.					
SQ	SEQUENCE 263 AA; 27821 MW; 2620A27AFBA5D04D CRC64;					
	Query Match	42.8%;	Score 71;	DB 11;	Length 263;	
	Best Local Similarity	48.3%;	Pred. No. 0.11;			
	Matches 14; Conservative	5;	Mismatches 10;	Indels 0;	Gaps	
QY	1 MLLSLTSLVLLGGSWGCGVPALTPALS 29	: ::: :				
Db	1 MAFLNLSCFALVGATFGCGVPALQPVL 29	: ::: :				
RESULT 6						
Q9DX8X	PRELIMINARY;		PRT;	263 AA.		
ID	Q9DX8X					
AC	Q9DX8X;					
DT	01-JUN-2001 (TREMBLrel. 17, Created)					
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)					
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)					
DE	2200008D09Rik protein.					
GN	2200008D09RIK.					
OS	Mus musculus (Mouse).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
OX	NCB1_TaxID=10090;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=C57BL/6J; TISSUE=PANCREAS;					
RX	MEDLINE=21085660; PubMed=11217851;					
RA	Kawai J., Shingagawa A., Shibata Y., Konno H., Adachi J., Fukuda S.,					
RA	Azakawa T., Hara A., Fukunishi Y., Kiyosava H., Kondo S., Yamataka I.,					
RA	Aizawa K., Izawa M., Nishi K., Kasakawa T., Saito R.,					
RA	Saito T., Okazaki Y., Gajobori T., Bono H., Sasavant T.,					
RA	Kadota K., Matsuda H.A., Ashburner T., Gissi C., King B., Kochiwa H.,					
RA	Fleischmann W., Gaasterland T., Nikaido I., Pesole G., Quackenbush J.,					
RA	Kuehl P., Lewis S., Matsuo F., Suzuki R., Tomita M., Wagner L., Washio T.,					
RA	Schram L.M., Staubli F., Furuno R., Anon H., Baldarelli R., Barsh G.,					
RA	Sakai K., Okido T., Bojunga N., Carninci P., de Bonaldo M.F.,					
RA	Blake J.J., Boffelli D., Fletcher C., Fujita M., Gariboldi M.,					
RA	Brownstein M.J., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,					
RA	Gustincich S., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,					
RA	Rondone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,					
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,					
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,					
RA	Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,					
RA	Hayashizaki Y.;					
RT	"Functional annotation of a full-length mouse cdna collection.";					
RL	Nature 409:685-690(2001).					
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE					
CC	TRYPSIN FAMILY.					
DR	EMBL; AK007566; BAB25112.1; -.					
DR	HSSP; P00766; 1GCT.					
DR	MEROPS; S01.152; -.					
DR	MGI;1913723; 2200008D09Rik.					
DR	InterPro: IPR001314; Chymotrypsin.					
DR	InterPro: IPR001254; Ser_protease_Try.					
DR	Pfam; PF00089; trypsin; 1.					
DR	PRINTS; PR00722; CHYMOTRYPSIN.					
DR	SMART; SM00020; TRYSP_SPC; 1.					

DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 263 AA; 27898 MW; C0638FB8F905A92F CRC64;

Query Match 42.8%; Score 71; DB 11; Length 263;
Best Local Similarity 48.3%; Pred. No. 0.11;
Matches 14; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGLSGWCGVPAITPALS 29
| | | | | : : : : : | | | | | : :
Db 1 MAFLWLVSFCALVGATFGGVPVPAIQPVLT 29

RESULT 7

ID Q9CR35 PRELIMINARY; PRT; 263 AA.

AC Q9CR35;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE 2200008D09R1k protein.
GN 2200008D09R1k.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=STOMACH, SPLEEN, AND PANCREAS;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa K., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE

TRYPsin FAMILY.

DR EMBL; AK008927; BAB25971.1; -

DR EMBL; AK003079; BAB22553.1; -

DR EMBL; AK007765; BAB25241.1; -

DR EMBL; AK007815; BAB25280.1; -

DR EMBL; AK008729; BAB25861.1; -

DR EMBL; AK008888; BAB25954.1; -

DR HSSP; P00766; LGCT.

DR MEROPS; S01.152; -

DR MGD; MGI:1913723; 2200008D09R1k.

DR InterPro; IPR001314; Chymotrypsin.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR SMART; SM00020; TRYP_SPC; 1.

DR PROSITE; PS00240; TRYPsin_DOM; 1.

DR PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.

DR PROSITE; PS00135; TRYPsin_SER; 1.

KW Hydrolase; Serine protease.

SQ SEQUENCE 263 AA; 27822 MW; 28C4487AF1A26B27 CRC64;

Query Match 42.8%; Score 71; DB 11; Length 263;
Best Local Similarity 48.3%; Pred. No. 0.11;
Matches 14; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGLSGWCGVPAITPALS 29
| | | | | : : : : : | | | | | : :
Db 1 MAFLWLVSFCALVGATFGGVPVPAIQPVLT 29

RESULT 8

ID Q9W7Q4

AC Q9W7Q4; PRELIMINARY; PRT; 261 AA.

DT 01-NOV-1999 (TREMBlrel. 12, Created)

DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Chymotrypsinogen 1.

OS Paralichthys olivaceus (Flounder).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;

OC Pleuronectidae; Paralichthyidae; Paralichthys.

OX NCBI_TaxID=8255;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=PANCREAS;

RA Suzuki T., Srivastava A.S., Kurokawa T.;

RT "Japanese flounder mRNA for chymotrypsinogen 1.";

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE

TRYPsin FAMILY.

DR EMBL; AB029753; BAA82365.1; -

DR HSSP; P00763; IDPO.

DR MEROPS; S01.256; -

DR InterPro; IPR001314; Chymotrypsin.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR SMART; SM00020; TRYP_SPC; 1.

DR PROSITE; PS00240; TRYPsin_DOM; 1.

DR PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.

DR PROSITE; PS00135; TRYPsin_SER; 1.

KW Hydrolase; Serine protease.

SQ SEQUENCE 261 AA; 28184 MW; D7090A9D65395B7D CRC64;

Query Match 35.8%; Score 59.5; DB 13; Length 261;
Best Local Similarity 48.5%; Pred. No. 3.8;
Matches 16; Conservative 5; Mismatches 9; Indels 3; Gaps 2;

QY 1 MLLSLTSLVLGLSGWCGVPAITPALS-YNQ 32

| | | | | : : : : : | | | | | : :
Db 1 MLLTSCFALV--ASALGCGVPSIKPQVSGYNK 31

RESULT 9

ID Q8THZ0

AC Q8THZ0; PRELIMINARY; PRT; 648 AA.

DT 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Hydrogenase-3, subunit C (HycC homolog).

GN MA4368.

OS Methanosarcina acetivorans.

OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;

OC Methanosarcinaceae; Methanosarcina.

OX NCBI_TaxID=2214;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C2A / ATCC 35395 / DSM 2834;

RX MEDLINE=21929760; PubMed=11932238;

RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,

```

RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., DeArellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Graham D.A., Guss A.M.,
RA Heiderich R., Ingram-Smith C., Kuetchner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.,
RA "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
KW EMBL; AE011156; AAM07110.1; -.
KW Complete proteome.
SQ SEQUENCE 648 AA; 70386 MW; 97CA36C7442E004B CRC64;

Query Match 35.5%; Score 59; DB 17; Length 648;
Best Local Similarity 22.4%; Pred. No. 11;
Matches 11; Conservative 8; Mismatches 12; Indels 18; Gaps 1;

Qy 1 MLLSLTSLVL-----LGSWGGCGVPAITPALSYN 31
:|:|:| | | : : :|:|:| | | :|:|:|
Db 504 LLLGLFLGLIYAALFVLTTRTKETNPISRTSETWRCGIPTVAPYMEYS 552

RESULT 10
Q96R18 PRELIMINARY; PRT; 216 AA.
ID Q96R18
AC Q96R18;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Olfactory receptor (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Fuchs T., Malecova B., Linhart C., Sharan R., Khen M., Herwig R.,
RA Shmulevich D., Eikon R., Steinfath M., O'Brien J.K., Radelof U.,
RA Lehach H., Olender Z., Gluesman G., Lancet D., Shamir R.;
RT "DFOG: A Practical Scheme for Deciphering Families of Genes.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF399627; AAK95112.1; -.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 1 216
FT NON_TER 216 216
SQ SEQUENCE 216 AA; 23670 MW; 57E6534FB0B22C75 CRC64;

Query Match 34.9%; Score 58; DB 4; Length 216;
Best Local Similarity 60.0%; Pred. No. 5;
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 2 LLLSLTSLVLGSSWGGCV 21
|:|:| :|:| | | | |
Db 67 LLLSHRVCLQAGSAWAGCV 86

RESULT 11
Q9W7Q3 PRELIMINARY; PRT; 260 AA.
ID Q9W7Q3
AC Q9W7Q3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Chymotrypsinogen 2.
OS Paralicthys olivaceus (Flounder).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidae; Paralicthidae; Paralicthys.
OX NCBI_TaxID=8255;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Suzuki T., Sivastava A.S., Kurokawa T.;
RT "Japanese flounder mRNA for chymotrypsinogen 2.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB029754; BAA82366.1; -.
DR HSP; P00766; ICHG.
DR MEROPS; S01.152; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 260 AA; 27793 MW; 9F583044E22F78C0 CRC64;

Query Match 33.7%; Score 56; DB 13; Length 260;
Best Local Similarity 45.0%; Pred. No. 11;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 10 LVLLGSSWGGCVPAITPALS 29
| :|:|:| | | | | | :|:|:|
Db 7 LAFAGAYGCGSPALPPVIT 26

RESULT 12
Q8SBN3 PRELIMINARY; PRT; 140 AA.
ID Q8SBN3
AC Q8SBN3;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Gp21 protein.
OS bacteriophage phi3626.
OC viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=190478;
RN [1]
RP SEQUENCE FROM N.A.
RA Zimmer M., Scherer S., Loessner M.J.;
RT "Bacteriophage phi3626 complete genome.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY082070; AAL96791.1; -.
SQ SEQUENCE 140 AA; 15944 MW; 8D395E570A0A153C CRC64;

Query Match 33.4%; Score 55.5; DB 9; Length 140;
Best Local Similarity 38.7%; Pred. No. 7.2;
Matches 12; Conservative 8; Mismatches 10; Indels 1; Gaps 1;

Qy 3 LLSLTSLVLGGS-SWGGCVPAITPALSYNQ 32
:|:|:| :|:|:| | | | | | | |
Db 4 ITAILSLVLFGLSGLLGGCKRETTNLINQ 34

RESULT 13
Q96666 PRELIMINARY; PRT; 441 AA.
ID Q96666
AC Q96666;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE IMAGINAL DISC growth factor 3.
GN IDGF3 OR CG4559.

```

OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kawamura K., Shibata T., Peel D., Park S.-B., Bryant P.J.;
 RT "A new family of growth factors produced by the fat body and active on
 RL Drosophila imaginal disc cells.";
 RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF102238; AAC99419.1; -;
 DR FlyBase: FBgn0020414; Idgf3.
 DR InterPro: IPR001223; Glyco_hydro_18.
 DR Pfam: PF00704; Glyco_hydro_18; 1.
 DR ProDom: PD000471; Glyco_hydro_18; 1.
 SQ SEQUENCE 441 AA; 49219 MW; 988C297D09CA0B08 CRC64;
 Query Match 33.4%; Score 55.5; DB 5; Length 441;
 Best Local Similarity 48.3%; Pred. No. 22;
 Matches 14; Conservative 5; Mismatches 9; Indels 1; Gaps 1;
 QY 2 LLLSLTSLVLLGSSWCGVPATPALSY 30
 Db 210 LLLSLTV-LPNVNSWYDAPSIAPSLDF 237
 RESULT 14
 Q9V3B9
 ID Q9V3B9 PRELIMINARY; PRT; 441 AA.
 AC Q9V3B9;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE IDGF3 protein (GH07453P).
 GN IDGF3 OR C64559.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA Sutton G.G., Wortman J.R., Randell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merklov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.M., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissensbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE-99403001; PubMed-10471707;
 RA Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R., Davis T.,
 RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
 RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
 RA Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,
 RA Celniker S., Rubin G.M.;
 RT "An exploration of the sequence of a 2.9-Mb region of the genome of
 RT Drosophila melanogaster: the Adh region.";
 RL Genetics 153:179-219(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA Celniker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G.,
 RA Butenhoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
 RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
 RA Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,
 RA Lewis S., Li P., Lomotan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
 RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
 RA Sethi H., Snir E., Svirskas R., Wan K.H., Weinburg T., Zhang R.,
 RA Zieran L.L., Rubin G.M.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuncio J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE003651; AAF53537.1; -;
 DR EMBL: AE003416; AAF45018.1; -;
 DR EMBL: AY047561; AAK77293.1; -;
 DR FlyBase: FBgn0020414; Idgf3.
 DR InterPro: IPR001223; Glyco_hydro_18.
 DR Pfam: PF00704; Glyco_hydro_18; 1.
 DR ProDom: PD000471; Glyco_hydro_18; 1.
 SQ SEQUENCE 441 AA; 49254 MW; F1938469FA117755 CRC64;
 Query Match 33.4%; Score 55.5; DB 5; Length 441;
 Best Local Similarity 48.3%; Pred. No. 22;
 Matches 14; Conservative 5; Mismatches 9; Indels 1; Gaps 1;
 QY 2 LLLSLTSLVLLGSSWCGVPATPALSY 30
 Db 210 LLLSLTV-LPNVNSWYDAPSIAPSLDF 237
 RESULT 15
 Q97Z55
 ID Q97Z55 PRELIMINARY; PRT; 835 AA.
 AC Q97Z55;
 DT 01-OCT-2001 (TREMBLrel. 18, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)

DE Hypothetical protein SSO1079.
GN SSO1079.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Awayez M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL; AE006726; AAK41341.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 835 AA; 92357 MW; BD28B5418BAF7250 CRC64;

Query Match 33.4%; Score 55.5; DB 17; Length 835;
Best Local Similarity 44.8%; Pred. NO. 40;
Matches 13; Conservative 5; Mismatches 4; Indels 7; Gaps 1;

QY 1 MLLSLTSLV-----LLGSSWGCGVP 22
:||||| : :||| |
Db 9 LLESLTLPPIPVNSQSTVVISWGNGTP 37

Search completed: December 20, 2002, 15:18:56
Job time : 6.08696 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:12:58 ; Search time 1.13043 Seconds
(without alignments)
1210.790 Million cell updates/sec

Title: US-09-856-319B-4_COPY_1_33

Perfect score: 166

Sequence: 1 MLLSLTSLVLLGSSGCGVPAITPALSYNQR 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	150	90.4	264	1	CTRL_HUMAN
2	76	45.8	263	1	CTRL_HUMAN
3	73	44.0	263	1	CTRL_HUMAN
4	67	40.4	263	1	CTRL_HUMAN
5	58	34.9	312	1	OAC1_HUMAN
6	55	33.1	232	1	FL3L_MOUSE
7	55	33.1	242	1	MTGA_KLEPN
8	55	33.1	268	1	CLCR_HUMAN
9	52	31.3	219	1	ISPD_CHLTR
10	51	30.7	484	1	DF5L_HUMAN
11	50	30.1	247	1	MCT5_MOUSE
12	50	30.1	363	1	CYB_TRYBB
13	50	30.1	742	1	KM65_YEAST
14	49.5	29.8	312	1	US10_HSV1
15	49	29.5	119	1	VTU3_DROME
16	49	29.5	222	1	CP19_SOLTU
17	49	29.5	344	1	YD33_MYCTU
18	48.5	29.2	65	1	MALX_KLEPN
19	48.5	29.2	310	1	PIP_SULSO
20	48.5	29.2	541	1	GT10_HUMAN
21	48.5	29.2	606	1	MM17_HUMAN
22	48	28.9	245	1	CTRA_BOVIN
23	48	28.9	245	1	CTRB_BOVIN
24	48	28.9	263	1	CTRA_GADMO
25	48	28.9	304	1	YG78_PSEAE
26	48	28.9	418	1	YK31_YEAST
27	48	28.9	650	1	DMWD_MOUSE
28	47.5	28.6	142	1	YOGA_BACSU
29	47.5	28.6	180	1	PAEP_HUMAN
30	47.5	28.6	218	1	ISPD_CHLMO
31	47	28.3	116	1	NU3M_STRCA
32	47	28.3	413	1	MOEA_SALTY
33	47	28.3	568	1	DNLI_METAC

34	47	28.3	832	1	BGAL_ASPOF
35	47	28.3	1620	1	ALK_HUMAN
36	46.5	28.0	118	1	NU3M_MARPO
37	46.5	28.0	255	1	THYL_RAT
38	46.5	28.0	262	1	NODJ_RHLIT
39	46.5	28.0	269	1	EL2_BOVIN
40	46	27.7	117	1	NU3M_SARGL
41	46	27.7	221	1	CP18_SOLTU
42	46	27.7	222	1	CP11_SOLTU
43	46	27.7	268	1	CLCR_RAT
44	46	27.7	324	1	NU1M_PISOC
45	46	27.7	365	1	TRBP_MOUSE

ALIGNMENTS

RESULT 1					
CTRL_HUMAN					
ID	CTRL_HUMAN	STANDARD;	PRT;	264	AA.
AC	P40313;				
DT	01-FEB-1995	(Rel. 31, Created)			
DT	01-FEB-1995	(Rel. 31, Last sequence update)			
DE	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Chymotrypsin-like protease CTRL-1 precursor (EC 3.4.21.-)				
GN	CTRL OR CTRL1				
OS	Homo sapiens (Human)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=94093544; PubMed=8268911;				
RA	Larsen F., Solheim J., Kristensen T., Kolsto A.B., Prydz H.;				
RT	"A tight cluster of five unrelated human genes on chromosome				
RL	16q22.1.";				
RL	Hum. Mol. Genet. 2:1589-1595(1993).				
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
EMBL	X71874; CAA50710.1; -				
EMBL	X71877; CAA50711.1; -				
HSSP	P00763; IDPO.				
MEROPS	S01.256; -				
GENE	HGNC:2524; CTRL.				
MIN	118888; -				
InterPro	IPR001314; Chymotrypsin.				
InterPro	IPR001254; Ser_protease_Try.				
Pfam	PF00089; trypsin; 1.				
PRINTS	PR00722; CHYMOTRYPSIN.				
SMART	SM00020; Tryp-SPC; 1.				
PROSITE	PS00240; TRYPsin_DOM; 1.				
PROSITE	PS00134; TRYPsin_HIS; 1.				
PROSITE	PS00135; TRYPsin_SER; 1.				
KW	Hydrolase; Serine protease; Glycoprotein; zymogen; Signal.				
SIGNAL	1 18				
PROPEP	19 33				
CHAIN	34 264				
ACT_SITE	75 75				
FT ACT_SITE	121 121				
FT ACT_SITE	214 214				
FT CARBOHYD	114 114				
FT DISULFID	19 141				
FT DISULFID	60 76				
FT DISULFID	155 220				
FT DISULFID	187 201				

34	47	28.3	832	1	BGAL_ASPOF
35	47	28.3	1620	1	ALK_HUMAN
36	46.5	28.0	118	1	NU3M_MARPO
37	46.5	28.0	255	1	THYL_RAT
38	46.5	28.0	262	1	NODJ_RHLIT
39	46.5	28.0	269	1	EL2_BOVIN
40	46	27.7	117	1	NU3M_SARGL
41	46	27.7	221	1	CP18_SOLTU
42	46	27.7	222	1	CP11_SOLTU
43	46	27.7	268	1	CLCR_RAT
44	46	27.7	324	1	NU1M_PISOC
45	46	27.7	365	1	TRBP_MOUSE

P45582	asparagus o
Q9um73	homo sapien
P26847	marchantia
P01150	rattus norv
P24144	rhizobium l
Q29461	bos taurus
Q63850	sarcophyton
O24384	solanum tub
P20347	solanum tub
P35091	rattus norv
P24995	pisaster oc
P97473	mus musculu

```

Query Match          45.8%; Score 76; DB 1; Length 263;
Best Local Similarity 55.2%; Pred. No. 0.0085;
Matches 16; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy  1  MLLSLTSLSLVLGSSWCGVPAITPALS 29
      | | | | | | | | | | | | | |
Db  1  NAFLWLLSCFALLGTAFGCGVPAIQPVLS 29

RESULT 3
CTRB_HUMAN          STANDARD;          PRT;   263 AA.
ID  * CTRB_HUMAN          AC
PI  P17538;
DT  01-AUG-1990 (Rel. 15, Created)
DT  01-AUG-1990 (Rel. 15, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Chymotrypsinogen B precursor (EC 3.4.21.1).
GN  CTRB1 OR CTRB.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Pancreas;
RX  MEDLINE=89134264; PubMed=2917002;
RA  Tomita N., Izumoto Y., Horii A., Doi S., Yokouchi H., Ogawa M.,
RA  Mori T., Matsubara K.;
RT  "Molecular cloning and nucleotide sequence of human pancreatic
RT  chymotrypsinogen cDNA.";
RL  Biochem. Biophys. Res. Commun. 158:569-575(1989).
[2]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Pancreas;
RA  Strausberg R.;
RL  Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC  -!- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
CC  Phe-I-Xaa, Leu-I-Xaa.
CC  -!- SUBCELLULAR LOCATION: Extracellular.
CC  -!- SIMILILARY: BELONGS TO PEPTIDASE FAMILY S1.
-----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
-----
EMBL; M24400; AAA52128.1; -.
DR  EMBL; BC005385; AA05385.1; -.
DR  PIR; A31299; A31299.
DR  HSP; P00766; ICHG.
DR  MEROPS; S01.152; -.
DR  Genew; HGNC:2521; CTRB1.
DR  MIM; 118890; -.
DR  InterPro; IPR001314; Chymotrypsin.
DR  InterPro; IPR001254; Ser.protease_Try.
DR  Pfam; PF00089; trypsin; 1.
DR  PRINTS; PR00722; CHYMOTRYPSIN.
DR  SMART; SM00020; Tryp_Spc; 1.
DR  PROSITE; PS00240; TRYPSIN_DOM; 1.
DR  PROSITE; PS00134; TRYPSIN_HIS; 1.
DR  PROSITE; PS00135; TRYPSIN_SER; 1.
KW  Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
FT  SIGNAL          1      18
FT  CHAIN           19     263      CHYMOTRYPSINOGEN B.
FT  CHAIN           19     31       CHYMOTRYPSIN B, A CHAIN.
FT  CHAIN           34     164       CHYMOTRYPSIN B, B CHAIN.
FT  CHAIN           167     263      CHYMOTRYPSIN B, C CHAIN.
FT  ACT_SITE        175     75       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT  ACT_SITE        120    120       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT  ACT_SITE        213    213       CHARGE RELAY SYSTEM (BY SIMILARITY).

```

FT DISULFID 19 140 BY SIMILARITY.
FT DISULFID 60 76 BY SIMILARITY.
FT DISULFID 154 219 BY SIMILARITY.
FT DISULFID 186 200 BY SIMILARITY.
FT DISULFID 209 238 BY SIMILARITY.
SQ SEQUENCE 263 AA; 27870 MW; 4C1C055A49088701 CRC64;

Query Match 44.0%; Score 73; DB 1; Length 263;
Best Local Similarity 55.2%; Pred. No. 0.021;
Matches 16; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLLGSSWCGVPAITPALS 29
| | | | | : : : : : | | | | |
DB 1 MAFWLSCWALLGTTCGVPPIHPVLS 29

RESULT 4
CTRB_RAT
ID CTRB_RAT STANDARD; PRT; 263 AA.
AC P07338;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsinogen B precursor (EC 3.4.21.1).
GN CTRB1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85054881; PubMed=6209274;
RA Bell G I., Quinto C., Quiroga M., Valenzuela P., Craik C.S.,
RA Rutter W.J.;
RT Isolation and sequence of a rat chymotrypsin B gene.";
RL J. Biol. Chem. 259:14265-14270(1984).
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
CC Phe-I-Xaa, Leu-I-Xaa.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; K02298; AAA98732.1; -;
DR PIR; A22658; KYRTP.
DR HSP; P00766; ICHG.
DR MEKOPS; S01.152; -;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser-protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp-SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
FT SIGNAL 1 18
FT CHAIN 19 263 CHYMOTRYPSINOGEN B.
FT CHAIN 19 31 CHYMOTRYPSIN B, A CHAIN.
FT CHAIN 34 164 CHYMOTRYPSIN B, B CHAIN.
FT CHAIN 167 263 CHYMOTRYPSIN B, C CHAIN.
FT ACT_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 19 140 BY SIMILARITY.
FT DISULFID 60 76 BY SIMILARITY.
FT DISULFID 154 219 BY SIMILARITY.

FT DISULFID 186 200 BY SIMILARITY.
FT DISULFID 209 238 BY SIMILARITY.
SQ SEQUENCE 263 AA; 27849 MW; ACAFD8AC8C8C64DA6D CRC64;
Query Match 40.4%; Score 67; DB 1; Length 263;
Best Local Similarity 44.8%; Pred. No. 0.12;
Matches 13; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLLGSSWCGVPAITPALS 29
| | | | | : : : : : | | | | |
DB 1 MAFWLVSFCALVGATFGCGVPTIQPVLT 29

RESULT 5
OACL_HUMAN
ID OACL_HUMAN STANDARD; PRT; 312 AA.
AC Q96KK4;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Olfactory receptor 10C1 (Hs6M1-17).
GN OR10C1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Whitaker H.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AL035542; CAB44507.1; ALT_INIT.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Multigene family; Olfaction.
FT DOMAIN 1 24 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 25 48 1 (POTENTIAL).
FT DOMAIN 49 56 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 57 78 2 (POTENTIAL).
FT DOMAIN 79 99 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 100 119 3 (POTENTIAL).
FT DOMAIN 120 138 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 139 157 4 (POTENTIAL).
FT DOMAIN 158 194 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 195 218 5 (POTENTIAL).
FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 236 258 6 (POTENTIAL).
FT DOMAIN 259 271 7 (POTENTIAL).
FT TRANSMEM 272 291 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 292 312 CYTOPLASMIC (POTENTIAL).
FT DISULFID 96 188 BY SIMILARITY.
FT CARBOHYD 4 4 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 312 AA; 34351 MW; B6FD6E0E700CB2CC CRC64;

Query Match 34.9%; Score 58; DB 1; Length 312;
Best Local Similarity 60.0%; Pred. No. 2.1;
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 LLLSLTSLVLLGSSWCGV 21

```
|||| : | | | | |
Db 133 LLLSHRVCLQAGSAMACGV 152

RESULT 6
FL3L_MOUSE STANDARD; PRT; 232 AA.
ID P49772; Q64085;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE SL cytokine precursor (Fms-related tyrosine kinase 3 ligand) (Flt3
ligand) (Flt3L).
GN FLT3LG OR FLT3L.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
RP MEDLINE=94195428; PubMed=8145851;
RA Hannum C., Culpepper J., Campbell D., McClanahan T., Zurawski S.,
BAZAN J.F., Kastelein R., Hudak S., Wagner J., Mattison J., Luh J.,
Duda G., Martina G., Peterson D., Menon S., Shanafelt A.,
Muench M., Kellner G., Nankawa R., Rennick D., Roncarolo M.G.,
Zlotnik A., Rosnet O., Dubreuil P., Birnbaum D., Lee F.;
*Ligand for FLT3/FLK2 receptor tyrosine kinase regulates growth of
RT haematopoietic stem cells and is encoded by variant RNAs.;
RL Nature 368:643-648(1994).
[2]
SEQUENCE FROM N.A.
RP STRAIN=5JL/J;
RA Lyman S.D., James L., Vandenbos T., Devries P., Brasel K.,
GLINKAK B., Hollingsworth L.T., Picha K.S., McKenna H.J.,
SPLETT R.R., Fletcher F.A., Maraskovsky E., Farrah T.,
Foxworth D., Williams D.E., Beckmann M.P.;
*Molecular cloning of a ligand for the flt3/flk-2 tyrosine kinase
RT receptor: a proliferative factor for primitive hematopoietic cells.;
RL Cell 75:1157-1167(1993).
[3]
SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
RP MEDLINE=96032581; PubMed=7566977;
RA Lyman S.D., Stocking K., Davidson B., Fletcher F., Johnson L.,
ESCOBAR S.;
*Structural analysis of human and murine flt3 ligand genomic loci.;
RL Oncogene 11:1165-1172(1995).
[4]
SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RP MEDLINE=95124710; PubMed=7824267;
RA Lyman S.D., James L., Escobar S., Downey H., de Vries P.,
BRASEL K., STOCKING K., BECKMANN M.P., COPELAND N.G.,
CLEVELAND L.S.;
*Identification of soluble and membrane-bound isoforms of the murine
RT flt3 ligand generated by alternative splicing of mRNAs.;
RL Oncogene 10:149-157(1995).
[5]
SEQUENCE FROM N.A.
RP McClanahan T., Culpepper J., Campbell D., Wagner J.,
FRANZ-BACON K., MATTISON J., TSAI S., LUH J., GUIMARES M.J.,
MATTEI M.-G., ROSNET O., BIRNBAUM D., HANNUM C.;
Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STIMULATES THE PROLIFERATION OF EARLY HEMATOPOIETIC
CC CELLS. SYNERGIZES WELL WITH A NUMBER OF OTHER COLONY STIMULATING
CC FACTORS AND INTERLEUKINS.
CC -1- SUBUNIT: Homodimer (soluble isoform) (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Two soluble
CC isoforms are also produced by alternative splicing. One of which,
CC isoform 3/E6, is biologically active, while the other, isoform
CC 4/E6delta16, is inactive.
CC -1- ALTERNATIVE PRODUCTS: 4 isoforms: 1/6C (shown here), 2/5H, 3/E6
CC and 4/E6delta16; are produced by alternative splicing.
-----

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; U04807; AAA18000.1; -
CC EMBL; L23636; AAA39436.1; -
CC EMBL; U29875; AAA90951.1; -
CC EMBL; U29875; AAA90952.1; -
CC EMBL; S76459; AAB33069.1; -
CC EMBL; S76461; AAB33070.1; -
CC EMBL; S76464; AAB33071.1; -
CC EMBL; U44024; AAA93307.1; -
CC EMBL; U44024; AAA93308.1; -
CC EMBL; MG1:95560; Flt3L.
CC InterPro: IPR004213; Flt3_lig.
CC Pfam: PF02947; flt3_lig; 1.
KW Cytokine; Glycoprotein; Transmembrane; Alternative splicing; Signal.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 232 SL CYTOKINE.
FT DOMAIN 190 210 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 211 232 POTENTIAL.
FT DOMAIN 31 112 CYTOPLASMIC (POTENTIAL).
FT DISULFID 71 156 BY SIMILARITY.
FT DISULFID 120 161 BY SIMILARITY.
FT DISULFID 127 127 BY SIMILARITY.
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 164 232 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 164 169 DSSTLLPRSPALAEATELPPEPRQLLLLLPLTLVL
FT VARSPLIC 170 232 LAAWGLRWQARRRGELHPGVLPSPH -> GNGGPRACH
FT VARSPLIC 159 163 HGATRLTATALLTVCPLLLPLVGTSHMFLPYFLSLSF
FT CONFLICT 141 141 LKMYLY (IN ISOFORM 2).
FT CONFLICT 198 198 DSSTLL -> GSHAG (IN ISOFORM 3).
FT CONFLICT 232 AA; 26141 MW; 3A3680D3CB69FBA6 CRC64; MISSING (IN REF. 2).
FT CONFLICT 141 141 A -> G (IN REF. 1).
FT CONFLICT 198 198 MISSING (IN REF. 2).
FT CONFLICT 232 AA; 26141 MW; 3A3680D3CB69FBA6 CRC64; MISSING (IN REF. 2).
SQ SEQUENCE 232 AA; 26141 MW; 3A3680D3CB69FBA6 CRC64; MISSING (IN REF. 2).

Query Match 33.1%; Score 55; DB 1; Length 232;
Best Local Similarity 61.1%; Pred. No. 3.8;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGSSWG 18
Db 193 LLLLLLPLTLVLLAAWG 210

RESULT 7
MTGA_KLEPN STANDARD; PRT; 242 AA.
ID MTGA_KLEPN
AC Q48465;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Monofunctional biosynthetic peptidoglycan transglycosylase
DE {EC 2.4.2.-} (Monofunctional Tgase).
GN MTGA.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=573;
[1]
SEQUENCE FROM N.A.
RP STRAIN=M5a1;
RX MEDLINE=96328711; PubMed=8830253;
RA Spratt B.G., Zhou J., Taylor M., Merrick M.J.;
RT "Monofunctional biosynthetic peptidoglycan transglycosylases.";
RL Mol. Microbiol. 19:639-640(1996).
```



```

OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UV-3/Cx;
RX MEDLINE=99000809; PubMed-9784136;
RA Stephens R.S., Kalman S., Lammell C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RL Chlamydia trachomatis";
RL Science 282:754-759(1998).
CC -!- FUNCTION: CATALYZES THE FORMATION OF 4-DIPHOSPHOCYTIDYL-2C-METHYL-
CC D-ERYTHRITOL FROM CTP AND 2C-METHYL-D-ERYTHRITOL 4-PHOSPHATE (BY
CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: CTP + 2-C-methyl-D-erythritol 4-phosphate =
CC diphosphate + 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol.
CC -!- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; third step.
CC -!- SIMILARITY: BELONGS TO THE ISPD FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001320; AAC68062.1; -
DR HSP: Q46893; 1.INI
DR InterPro: IPR001228; ISPD_synthase.
DR Pfam: PF01128; ISPD; 1.
DR PROSITE: PS01295; ISPD; 1.
KW Transferase; Nucleotidyltransferase; Isoprene biosynthesis;
KW Complete proteome.
SQ SEQUENCE 219 AA; 24168 MW; 45C69BF82505364D CRC64;

Query Match 31.3%; Score 52; DB 1; Length 219;
Best Local Similarity 36.4%; Pred. No. 8.8;
Matches 16; Conservative 3; Mismatches 11; Indels 14; Gaps 1;

QY 4 LSITSLVLGGSGW-----CGVPAITPALSYNQR 33'
DB 1 MNLSCSLVLGGGKGRFNSLPKQYTHLCGPELILHALHAYOR 44

RESULT 10
DF5L_HUMAN
ID DF5L_HUMAN STANDARD; PRT; 484 AA.
AC P57764;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DFNS-like protein FLJ12150.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saico K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;

```

```

RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE DFNS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AK022212; BAB13986.1; -
DR EMBL: BC008904; AAH08904.1; -
SQ SEQUENCE 484 AA; 52801 MW; F7CE8073E0C0194D CRC64;

Query Match 30.7%; Score 51; DB 1; Length 484;
Best Local Similarity 63.2%; Pred. No. 24;
Matches 12; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

QY 7 TLSLV--LLGSSWGGCVPA 23
DB 420 TWSLPPGLLGSWGEGAPA 438

RESULT 11
MCT5_MOUSE
ID MCT5_MOUSE STANDARD; PRT; 247 AA.
AC P21844;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mast cell protease 5 precursor (EC 3.4.21.-) (MMCP-5) (Mast cell
DE chymase 1).
GN MCP5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92041862; PubMed-1939089;
RA McNeill H.P., Austen K.F., Somerville L.L., Gurish M.F., Stevens R.L.;
RT "Molecular cloning of the mouse mast cell protease-5 gene. A novel
RT secretory granule protease expressed early in the differentiation of
RT serosal mast cells.";
RL J. Biol. Chem. 266:20316-20322(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Leadon X A1; TISSUE=Connective tissue;
RX MEDLINE=91285010; PubMed-2060576;
RA Huang R., Blom T., Hellman L.;
RT "Cloning and structural analysis of MMCP-1, MMCP-4 and MMCP-5, three
RT mouse mast cell-specific serine proteases.";
RL Eur. J. Immunol. 21:1611-1621(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92287966; PubMed-1376147;
RA Chu W., Johnson D.A., Musich P.R.;
RT "Molecular cloning and characterization of mouse mast cell chymases.";
RL Biochim. Biophys. Acta 1121:83-87(1992).
RN [4]
RP SEQUENCE OF 22-51.
RX MEDLINE=90222202; PubMed-2326280;
RA Reynolds D.S., Stevens R.L., Lane W.S., Carr M.H., Austen K.F.,
RA Serafini W.E.;
RT "Different mouse mast cell populations express various combinations
RT of at least six distinct mast cell serine proteases.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:3230-3234(1990).
RN [5]
RP SUBCELLULAR LOCATION
RX MEDLINE=92407343; PubMed-1527387;
RA McNeill H.P., Frenkel D.P., Austen F., Friend D.S., Stevens R.L.;
RT "Translation and granule localization of mouse mast cell protease-5.

```


Immunodetection with specific antipeptide Ig.";
 J. Immunol. 149:2466-2472(1992).
 -!- SUBCELLULAR LOCATION: Secretory granules.
 -!- TISSUE SPECIFICITY: MAST CELLS.
 -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. GRANYME SUBFAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL: X68805; CAA48705.1; ALT_INIT.
 EMBL: M73759; AAM40105.1; -.
 EMBL: M73760; -; NOT_ANNOTATED_CDS.
 EMBL: M68898; AAA39492.1; -.
 PIR: A41076; A41076.
 DR: C35646; C35646.
 HSSP: P23946; 1KLT.
 MEROPS: S01.150; -.
 MGD: MGI:96941; Mcpt5.
 InterPro: IPR001314; Chymotrypsin.
 InterPro: IPR001254; Ser_protease_Try.
 Pfam: PF00089; trypsin; 1.
 PRINTS: PR00722; CHYMOTRYPSIN.
 SMART: SM00020; Tryp_SPC; 1.
 PROSITE: PS02040; TRYPSIN_DOM; 1.
 PROSITE: PS00134; TRYPSIN_HIS; 1.
 PROSITE: PS00135; TRYPSIN_SER; 1.
 KW Hydrolyase; Serine protease; Signal; Glycoprotein; Zymogen.
 FT SIGNAL 1 19
 FT PROPEP 20 21 POTENTIAL.
 FT CHAIN 22 247 ACTIVATION PEPTIDE.
 FT ACT_SITE 66 66 MAST CELL PROTEASE 5.
 FT ACT_SITE 110 110 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 203 203 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 51 67 BY SIMILARITY.
 FT DISULFID 144 209 BY SIMILARITY.
 FT DISULFID 175 188 BY SIMILARITY.
 FT CARBOHYD 80 80 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLICT 5 5 T -> A (IN REF. 1).
 FT CONFLICT 51 51 C -> R (IN REF. 4).
 FT CONFLICT 224 224 A -> R (IN REF. 3).
 SQ SEQUENCE 247 AA; 27586 MW; 24C290CF61237DC7 CRC64;

 Query Match 30.1%; Score 50; DB 1; Length 247;
 Best Local Similarity 39.5%; Pred. No. 18;
 Matches 15; Conservative 5; Mismatches 10; Indels 8; Gaps 1;
 OY 1 MLLSLTLVLGSSWCGG-----VPAITPALS 30
 DB 1 MHLTLHLGSSSTKAGIEIGTECIPHSRPMAY 38

 RESULT 12
 CVB_TRYBB
 ID CVB_TRYBB STANDARD; PRT; 363 AA.
 AC P00164;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Cytochrome B.
 GN COB OR CYTB.
 OS Trypanosoma brucei brucei.
 OG Mitochondrion.
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 CC NCBI_TaxID=5702;
 RN [1]
 RP SEQUENCE FROM N.A. (KINETOPLAST).
 RX MEDLINE=84041494; PubMed=6314266;
 RA Benne R., de Vries B.F., van den Burg J., Klaver B.;

"The nucleotide sequence of a segment of Trypanosoma brucei
 mitochondrial maxi-circle DNA that contains the gene for
 apocytocrome b and some unusual unassigned reading frames.";
 Nucleic Acids Res. 11:6925-6941(1983).
 -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
 COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
 RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 COUPLED TO ATP SYNTHESIS.
 -!- COPACITOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 BOUND TO THE PROTEIN.
 -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CYTOCHROME C1 AND THE RIESKE PROTEIN.
 -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL: X00017; CAA24915.1; -.
 PIR: A00160; CBUTB.
 InterPro: IPR000179; Cyt_b_b6.
 InterPro: IPR001230; Prenyl_site.
 Pfam: PF00032; cytochrome_b_c1; 1.
 Pfam: PF00033; cytochrome_b_n; 1.
 PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
 PROSITE: PS00193; CYTOCHROME_B_OO; 1.
 KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
 Heme; Kinetoplast.
 FT METAL 74 74 IRON 1 (HEME B562 AXIAL LIGAND).
 FT METAL 88 88 IRON 2 (HEME B566 AXIAL LIGAND).
 FT METAL 175 175 IRON 2 (HEME B562 AXIAL LIGAND).
 FT METAL 189 189 IRON 1 (HEME B566 AXIAL LIGAND).
 SQ SEQUENCE 363 AA; 43455 MW; 2A37CD827E55923B CRC64;

 Query Match 30.1%; Score 50; DB 1; Length 363;
 Best Local Similarity 46.4%; Pred. No. 25;
 Matches 13; Conservative 3; Mismatches 12; Indels 0; Gaps 0;
 OY 3 LLSLTLVLGSSWCGGVPVPAITPALS 30
 DB 314 LMLTYSLLLFYSINWSGFLALYVVLAY 341

 RESULT 13
 KM65_YEAST
 ID KM65_YEAST STANDARD; PRT; 742 AA.
 AC O03656;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable serine/threonine-protein kinase YMR216C (BC 2.7.1.-).
 GN YMR216C OR YMR261.10C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 CC NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Dedman K., Brown D., Bowman S., Barrell B.G., Rajandream M.A.,
 Walsh S.V.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC SIMILAR TO S.POMBE DSK1.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: Z49809; CAA89931.1; -
DR SGD: S0004829; YMR216C.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 2.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
KW ATP-binding.
FT DOMAIN 158 706 PROTEIN KINASE.
FT NP_BIND 164 172 ATP (BY SIMILARITY).
FT BINDING 187 187 ATP (BY SIMILARITY).
FT ACT_SITE 294 294 BY SIMILARITY.
SQ SEQUENCE 742 AA; 83237 MW; C775F10B30C950FC CRC64;

Query Match 30.1%; Score 50; DB 1; Length 742;
Best Local Similarity 80.0%; Pred. No. 49;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 VLLGSSWGCG 20
ID VLLGSSWGCG 20
DB 575 VLLGAPWGCG 584

RESULT 14
US10_HSV11
ID US10_HSV11 STANDARD; PRT; 312 AA.
AC P06486;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Virion protein US10.
GN US10.
OS Herpes simplex virus (type 1 / strain 17).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85160822; PubMed=2984429;
RA McGeoch D.J., Dolan A., Donald S., Rixon F.J.;
RT "Sequence determination and genetic content of the short unique
RL region in the genome of herpes simplex virus type 1.";
RL J. Mol. Biol. 181:1-13(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=84169548; PubMed=6324121;
RA Rixon F.J., McGeoch D.J.;
RT "A 3' co-terminal family of mRNAs from the herpes simplex virus type
RT 1 short region: two overlapping reading frames encode unrelated
RT polypeptide one of which has highly reiterated amino acid sequence.";
RL Nucleic Acids Res. 12:2473-2487(1984).
CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUP TOGETHER HSV-1 US10,
CC EHV-1 66, EHV-4 ORF3, AND VZV 64/69.
CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X14112; CAA32275.1; -
DR EMBL: X02138; CAA26064.2; -
DR EMBL: X00428; CAA25126.1; -
DR EMBL: L00036; AAA96678.1; -

DR PIR: A05242; QOBE07.
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR000714; EHV_Unk.
DR Pfam: PF02053; Gene66; 1.
KW Zinc-finger. 271 293 POTENTIAL.
FT 2N_FING
SQ SEQUENCE 312 AA; 34055 MW; 12D01B0E7C920EA3 CRC64;

Query Match 29.8%; Score 49.5; DB 1; Length 312;
Best Local Similarity 54.5%; Pred. No. 26;
Matches 12; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

QY 7 TSLVLGGSSWGCGVPAITPAL 28
ID TSLVLGGSSWGCGVPAITPAL 28
DB 291 THCLHLFG-AFGGDPALTPPL 311

RESULT 15
VTU3_DROME
ID VTU3_DROME STANDARD; PRT; 119 AA.
AC Q06521;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Vitelline membrane protein Vm34Ca precursor.
GN Vm34CA OR Vm34C.1.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94123988; PubMed=8293994;
RA Scherer L.J., Harris D.H., White M.K., Steel L.S., Jin J.,
RA Petri W.H.;
RT "Comparative analysis of the sequence and structure of two Drosophila
RL melanogaster genes encoding vitelline membrane proteins.";
RL Gene 136:121-127(1993).
RN [2]
RP SEQUENCE OF 1-96 FROM N.A.
RC STRAIN=Daekwanryeong;
RX MEDLINE=85257433; PubMed=3926479;
RA Mindrinos M.N., Scherer L.J., Garcini F.J., Kwan H., Jacobs K.A.,
RA Petri W.H.;
RT "Isolation and chromosomal location of putative vitelline membrane
RT genes in Drosophila melanogaster.";
RL EMBO J. 4:147-153(1985).
RN [3]
RP SEQUENCE OF 69-106 FROM N.A.
RC STRAIN=Oregon-R;
RX MEDLINE=89065296; PubMed=3143615;
RA Scherer L.J., Harris D.H., Petri W.H.;
RT "Drosophila vitelline membrane genes contain a 114 base pair region
RT of highly conserved coding sequence.";
RL Dev. Biol. 130:786-788(1988).
CC -1- FUNCTION: MAJOR EARLY EGG-SHELL PROTEIN.
CC -1- TISSUE SPECIFICITY: FOLLICLE CELLS.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING VITELLINE MEMBRANE
CC BIOSYNTHESIS.
CC -1- SIMILARITY: A 38 AMINO ACIDS REGION (VM DOMAIN) IS CONSERVED IN
CC DROSOPHILA VITELLINE MEMBRANE PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: L08852; AAC37200.1; -
DR EMBL: X01802; CAA25933.1; -

DR EMBL; M22700; AAA29018.1; --
 KW FlyBase; FBgn0003983; Vm34Ca.
 FT SIGNAL; Structural protein; Eggshell.
 FT SIGNAL 1 19
 FT CHAIN 20 119
 FT DOMAIN 69 106
 FT CONFLICT 96 96
 FT CONFLICT 96 96
 SQ SEQUENCE 119 AA; 11934 MW; 2F06298E52005BFC CRC64;

Query Match 29.5%; Score 49; DB 1; Length 119;
 Best Local Similarity 54.5%; Pred. No. 12;
 Matches 12; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

OY 12 LLGSSW--GCGVPAITPALSYN 31
 :||||: | ||| || ||:
 Db 25 MLGSSYGGYGKPAAPAPSYS 46

Search completed: December 20, 2002, 15:17:14
 Job time : 2.13043 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:16:49 ; Search time 8.34783 Seconds
(without alignments)
530.147 Million cell updates/sec

Title: US-09-856-319B-4
Perfect score: 1418
Sequence: 1 MLLSLTLTLVLLSSWGGC.....AMYTRVSKFSTWVQVWYN 264

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 106657 seqs, 16763532 residues

Total number of hits satisfying chosen parameters: 106657

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	889	62.7	192	10	US-09-925-297-529
2	822.5	58.0	263	10	US-09-888-615-96
3	485.5	34.2	812	9	US-09-335-325-1
4	485.5	34.2	812	10	US-09-788-142-1
5	485.5	34.2	812	10	US-09-761-120-1
6	485.5	34.2	812	10	US-09-873-676-81
7	484.5	34.2	270	10	US-09-923-779-152
8	484.5	34.2	273	10	US-09-925-297-695
9	484	34.1	791	9	US-09-967-386-1
10	484	34.1	810	10	US-09-946-893-2
11	468.5	33.0	453	9	US-09-978-295A-69
12	468.5	33.0	453	9	US-09-978-697-69
13	468.5	33.0	453	9	US-09-978-192A-69
14	468.5	33.0	453	12	US-10-052-586-64
15	468	33.0	248	10	US-09-925-301-1017
16	468	33.0	327	10	US-09-804-156-16
17	468	33.0	327	10	US-09-946-633-8
18	468	33.0	454	10	US-09-888-615-103
19	464.5	32.8	251	9	US-09-961-721-4

20	462.5	32.6	343	10	US-09-948-094-2	Sequence 2, Appli
21	457.5	32.3	249	9	US-09-961-721-5	Sequence 5, Appli
22	443.5	31.3	226	9	US-09-961-721-6	Sequence 6, Appli
23	443.5	31.3	226	9	US-10-045-367A-5	Sequence 5, Appli
24	443.5	31.3	226	10	US-09-910-151-6	Sequence 6, Appli
25	440.5	31.1	1169	9	US-09-870-759-126	Sequence 126, App
26	440	31.0	393	9	US-10-012-896-934	Sequence 934, App
27	440	31.0	393	10	US-09-759-143-934	Sequence 934, App
28	440	31.0	393	10	US-09-780-669-934	Sequence 934, App
29	440	31.0	393	10	US-09-822-827-934	Sequence 934, App
30	440	31.0	492	9	US-10-012-896-932	Sequence 932, App
31	440	31.0	492	10	US-09-759-143-932	Sequence 932, App
32	440	31.0	492	10	US-09-780-669-932	Sequence 932, App
33	440	31.0	492	10	US-09-822-827-932	Sequence 932, App
34	438	30.9	492	9	US-10-012-896-895	Sequence 895, App
35	438	30.9	492	10	US-09-759-143-895	Sequence 895, App
36	438	30.9	492	10	US-09-780-669-895	Sequence 895, App
37	438	30.9	492	10	US-09-879-792-14	Sequence 14, Appl
38	438	30.9	492	10	US-09-822-827-895	Sequence 895, App
39	437	30.8	384	9	US-09-981-353-23	Sequence 23, Appl
40	435.5	30.7	970	10	US-09-888-615-101	Sequence 101, App
41	434.5	30.6	320	10	US-09-888-615-90	Sequence 90, Appl
42	431.5	30.4	284	12	US-10-041-400A-7	Sequence 7, Appli
43	431.5	30.4	284	12	US-10-041-264A-7	Sequence 7, Appli
44	431.5	30.4	284	12	US-10-042-091A-7	Sequence 7, Appli
45	431	30.4	283	10	US-09-988-975A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-925-297-529
; Sequence 529, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 529
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-297-529

Query Match	62.7%	Score 889;	DB 10;	Length 192;
Best Local Similarity	84.7%	Pred. No. 1e-71;		
Matches 160;	Conservative 19;	Mismatches 10;	Indels 0;	Gaps 0;
QY	5	SUTLSLVLGSSWCCGVPATPALSYNQIRVINGENAVPGSWPQVSLQDNTGFHFCGSL	64	
Db	1	SUTLSLVLGSSWCCGIPAKPALSFSQIRVINGENAVLGSPWPQVSLQDSSGFHFCGSL	60	
QY	65	ISPNVVVTAHCOVTPGRHFVVLGEYDRSSNAEPVQVLSARAITHPNNANMTMNDLTL	124	
Db	61	ISQSVVVVTAHCVNVPGRHFVVLGEYDRSSNAEPQVLSVSRAITHPNSWNSTMMNDVTL	120	
QY	125	LKLASPARVTAQVSPVCLASTNEALPSGLTCVTTGGRISGVGNVTPARLQVVLPLVTV	184	
Db	121	LKLASPAQVTRISPVCLASSNEALTEGLTCVTTGGRISGVGNVTPAHLQVVALPLVTV	180	
QY	185	NOCROYWGA	193	
Db	181	NOCROYWGS	189	

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/788.142
FILING DATE: 16-Feb-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/866,735
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Warren, William L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 05940-0129

TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 818-3700
TELEFAX: (404) 818-3799

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 812 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

ORIGINAL SOURCE:

ORGANISM: Murine

IMMEDIATE SOURCE:

CLONE: Plasminogen

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-788-142-1

Query Match 34.2%; Score 485.5; DB 10; Length 812;
Best Local Similarity 40.0%; Pred. No. 2.9e-35;
Matches 106; Conservative 35; Mismatches 107; Indels 17; Gaps 7;

Qy 8 LSLVLLGSSWCGVPAITPALSYNQRIYNGENAVPGSWPQVSLQDN-TGFHFCGSLIS 66
Db 557 IPLCASASSFCGKQVPEPK-KCGRVVGCVANPHSPWQISLRTFTGQHFCCGTLIA 615
Qy 67 PNWVYTAHC---QVTPGRHFVVLGEYDRSSNAEPVQVLSIARATHPNWNANTMNDLT 123
Db 616 PEWLTAAHCLKSSRPEFYKVLGAHEEYIRGLDVQEISVAKLILEPN-----NRDIA 669
Qy 124 LLKLASPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQOVVPLVT 183
Db 670 LLKLSRPATITDKVIPACLPSPNYMVDRTICYITGWTGTGFGA--GRLEAQLPVIE 727
Qy 184 VNQCR--QYWGARTIDAMICAG--GSGASSCGDGGPLVCOKGNTWVLIGIVSWGKNC 239
Db 728 NKVCNRVEYLNRRVKSTELCAGLAGVDSCQDGGPLVCPEKDKYILOGVTSMGLGCA 787
Qy 240 NIQAPAMTRYSKFSTWINQVMAYN 264
Db 788 RPNKPGVYVRVSRFVDWIEREMRN 812

RESULT 5
US-09-761-120-1

Sequence 1, Application US/09761120
Patent No. US2002037847A1
GENERAL INFORMATION:
APPLICANT: Folkman, M. Judah
APPLICANT: O'Reilly, Michael

TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasminogen
FILE REFERENCE: 05940-0151 (43171-252068)
CURRENT APPLICATION NUMBER: US/09/761.120
CURRENT FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 09/309,821

PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 08/866,735
PRIOR FILING DATE: 1997-05-30
NUMBER OF SEQ ID NOS: 47
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 812
TYPE: PRT
ORGANISM: Murinae gen. sp.
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Plasminogen
US-09-761-120-1

Query Match 34.2%; Score 485.5; DB 10; Length 812;
Best Local Similarity 40.0%; Pred. No. 2.9e-35;
Matches 106; Conservative 35; Mismatches 107; Indels 17; Gaps 7;

Qy 8 LSLVLLGSSWCGVPAITPALSYNQRIYNGENAVPGSWPQVSLQDN-TGFHFCGSLIS 66
Db 557 IPLCASASSFCGKQVPEPK-KCGRVVGCVANPHSPWQISLRTFTGQHFCCGTLIA 615
Qy 67 PNWVYTAHC---QVTPGRHFVVLGEYDRSSNAEPVQVLSIARATHPNWNANTMNDLT 123
Db 616 PEWLTAAHCLKSSRPEFYKVLGAHEEYIRGLDVQEISVAKLILEPN-----NRDIA 669
Qy 124 LLKLASPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQOVVPLVT 183
Db 670 LLKLSRPATITDKVIPACLPSPNYMVDRTICYITGWTGTGFGA--GRLEAQLPVIE 727
Qy 184 VNQCR--QYWGARTIDAMICAG--GSGASSCGDGGPLVCOKGNTWVLIGIVSWGKNC 239
Db 728 NKVCNRVEYLNRRVKSTELCAGLAGVDSCQDGGPLVCPEKDKYILOGVTSMGLGCA 787
Qy 240 NIQAPAMTRYSKFSTWINQVMAYN 264
Db 788 RPNKPGVYVRVSRFVDWIEREMRN 812

RESULT 6
US-09-873-676-81

Sequence 81, Application US/09873676
Patent No. US20020077289A1
GENERAL INFORMATION:
APPLICANT: Macdonald, Nicholas J.

TITLE OF INVENTION: Angiotensin and Endostatin Binding Proteins and Methods of Use
FILE REFERENCE: 05213-0378 (43170-259333)
CURRENT APPLICATION NUMBER: US/09/873.676
CURRENT FILING DATE: 2001-06-04

PRIOR APPLICATION NUMBER: US 60/209,065
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: US 60/289,387
PRIOR FILING DATE: 2001-05-08
NUMBER OF SEQ ID NOS: 123
SOFTWARE: PatentIn version 3.1
SEQ ID NO 81
LENGTH: 812
TYPE: PRT
ORGANISM: Murinae gen. sp.
US-09-873-676-81

Query Match 34.2%; Score 485.5; DB 10; Length 812;
Best Local Similarity 40.0%; Pred. No. 2.9e-35;
Matches 106; Conservative 35; Mismatches 107; Indels 17; Gaps 7;

Qy 8 LSLVLLGSSWCGVPAITPALSYNQRIYNGENAVPGSWPQVSLQDN-TGFHFCGSLIS 66
Db 557 IPLCASASSFCGKQVPEPK-KCGRVVGCVANPHSPWQISLRTFTGQHFCCGTLIA 615
Qy 67 PNWVYTAHC---QVTPGRHFVVLGEYDRSSNAEPVQVLSIARATHPNWNANTMNDLT 123
Db 616 PEWLTAAHCLKSSRPEFYKVLGAHEEYIRGLDVQEISVAKLILEPN-----NRDIA 669

Db	545	SFCGKPKQVEPK-KCPRVVGVCVAPHSPWQVSLTRFGMHFCGGTILISPEWLTAAH	603
QY	76	C---QVTPGRHFVVLGEYDRSSNAEP-VQVLSIARATHIPNNNANTMNDLTLLKLASPA	131
Db	604	CLEKSPRPSSYKYLGAH-QEVNLEPHVQIEVSRLEPT-----RNDIALLLKSSPA	656
QY	132	RYTAQVSPVCLASTNEALPSGLTCVTTGRTSGVGNVTTPARLQVVLPTVTVNQCQRY-	190
Db	657	VITDKVIPACLPSPNTVVAADRTCEFTIGWGETGTGFGA--GLLKEAQLPVIEKNVCNRYE	714
QY	191	-WGARTDMAMICAG--GSAGSSCGDGGPLVCQKGNWTVLIGIVSWGKNCNIQAPAY	247
Db	715	FLNGRVQSTELCAGHLAGGTDCQDSSGGLPVCFEKDKYILQVTSWGLGCARPKNPGVI	774
QY	248	TRVSKFTSTWVQVMAYN	264
Db	775	VRVSREFTVWIEGVMRNN	791

```

RESULT 10
US-09-946-893-2
; Sequence 2, Application US/09946893
; Patent No. US20020072494A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yi-hai
; TITLE OF INVENTION: Materials and methods relating to endothelial cell growth
; FILE REFERENCE: Inhibitors
; FILE REFERENCE: Newburn
; CURRENT APPLICATION NUMBER: US/09/946,893
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/230,893
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 810
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-946-893-2

```

RESULT 11
US-978-295A-69
; Sequence 69, Application US/09978295A
; Patent No. US2002015606A1
; GENERAL INFORMATION:
; APPLICANT: ASHKENAZI, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David

1	APPLICANT:	Desnoyers, Luc	
2	APPLICANT:	Eaton, Dan	
3	APPLICANT:	Ferrara, Napoleon	
4	APPLICANT:	Filvaroff, Ellen	
5	APPLICANT:	Fong, Sherman	
6	APPLICANT:	Gao, Wei-Qiang	
7	APPLICANT:	Gerber, Hanspeter	
8	APPLICANT:	Gerritsen, Mary E.	
9	APPLICANT:	Goddard, Audrey	
10	APPLICANT:	Godowski, Paul J.	
11	APPLICANT:	Grimaldi, J. Christopher	
12	APPLICANT:	Gurney, Austin L.	
13	APPLICANT:	Hillan, Kenneth J.	
14	APPLICANT:	Kljavin, Ivar J.	
15	APPLICANT:	Kuo, Sophia S.	
16	APPLICANT:	Napier, Mary A.	
17	APPLICANT:	Pan, James;	
18	APPLICANT:	Paoni, Nicholas F.	
19	APPLICANT:	Roy, Margaret Ann	
20	APPLICANT:	Shelton, David L.	
21	APPLICANT:	Stewart, Timothy A.	
22	APPLICANT:	Tumas, Daniel	
23	APPLICANT:	Williams, P. Mickey	
24	APPLICANT:	Wood, William I.	
25	TITLE OF INVENTION:	Secreted and Transmembrane Polypeptides and Nucleic	
26	TITLE OF INVENTION:	Acids Encoding the Same	
27	FILE REFERENCE:	P2630Pic11	
28	CURRENT APPLICATION NUMBER:	US/09/978,295A	
29	CURRENT FILING DATE:	2001-10-15	
30	PRIOR APPLICATION NUMBER:	09/918585	
31	PRIOR FILING DATE:	2001-07-30	
32	PRIOR APPLICATION NUMBER:	60/062250	
33	PRIOR FILING DATE:	1997-10-17	
34	PRIOR APPLICATION NUMBER:	60/064249	
35	PRIOR FILING DATE:	1997-11-03	
36	PRIOR APPLICATION NUMBER:	60/065311	
37	PRIOR FILING DATE:	1997-11-13	
38	PRIOR APPLICATION NUMBER:	60/066364	
39	PRIOR FILING DATE:	1997-11-21	
40	PRIOR APPLICATION NUMBER:	60/077450	
41	PRIOR FILING DATE:	1998-03-10	
42	PRIOR APPLICATION NUMBER:	60/077632	
43	PRIOR FILING DATE:	1998-03-11	
44	PRIOR APPLICATION NUMBER:	60/077641	
45	PRIOR FILING DATE:	1998-03-11	
46	PRIOR APPLICATION NUMBER:	60/077649	
47	PRIOR FILING DATE:	1998-03-11	
48	PRIOR APPLICATION NUMBER:	60/077791	
49	PRIOR FILING DATE:	1998-03-12	
50	PRIOR APPLICATION NUMBER:	60/078004	
51	PRIOR FILING DATE:	1998-03-13	
52	PRIOR APPLICATION NUMBER:	60/078886	
53	PRIOR FILING DATE:	1998-03-20	
54	PRIOR APPLICATION NUMBER:	60/078936	
55	PRIOR FILING DATE:	1998-03-20	
56	PRIOR APPLICATION NUMBER:	60/078910	
57	PRIOR FILING DATE:	1998-03-20	
58	PRIOR APPLICATION NUMBER:	60/078939	
59	PRIOR FILING DATE:	1998-03-20	
60	PRIOR APPLICATION NUMBER:	60/079294	
61	PRIOR FILING DATE:	1998-03-25	
62	PRIOR APPLICATION NUMBER:	60/079656	
63	PRIOR FILING DATE:	1998-03-26	
64	PRIOR APPLICATION NUMBER:	60/079664	
65	PRIOR FILING DATE:	1998-03-27	
66	PRIOR APPLICATION NUMBER:	60/079689	
67	PRIOR FILING DATE:	1998-03-27	
68	PRIOR APPLICATION NUMBER:	60/079663	
69	PRIOR FILING DATE:	1998-03-27	
70	PRIOR APPLICATION NUMBER:	60/079728	
71	PRIOR FILING DATE:	1998-03-27	
72	PRIOR APPLICATION NUMBER:	60/079786	
73	PRIOR FILING DATE:	1998-03-27	

;
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558

;
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 33.0%; Score 468.5; DB 9; Length 453;

Best Local Similarity 38.1%; Pred. No. 4.5e-34;

Matches 91; Conservative 40; Mismatches 97; Indels 11; Gaps 6;

QY 30 YNORIVGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVTAHNC---QVTPGRHFVV 86
Db 213 YSSRIYGGNNLSLQWPAQSLQFQ-GYHLCGGSVITPLWIIITAAHCYDLYLPKSWTIQ 271
QY 87 LGEYDRSSNAEPYQVLSIARAITHPNNANTMNDLTLKLSAPRYTAQVSPVCLASTN 146
Db 272 VGLVSLLDNPPASHL--VEKIVYHSKYKPRKLGNDLALMKLAGPLTFENEMIQVCLPNS 329
QY 147 EALPSGLTCVTTGWRISGVNVTTPARLQOVLPVLTVNOCC--ROYWGARITDAMICAG- 203
Db 330 ENFPDGKVCWTSCWATEDGGDASPV-LNHAAVPLISNKNICNHRDVGKGIISPSMLCAGY 388
QY 204 -GSCASSCQDSDGGLVPCQKQNTWVLIGIVSWGCTKNCNIOAPAMYTRVSKFSTWLNQVM 261
Db 389 LTGGVDSQCGDSGGPLVCQERRLKLVGATSGFGICAEVNKPGVYTRVTSFLDWIHEQM 447

RESULT 12

US-09-978-697-69

; Sequence 69, Application US/09978697

Patent No. US20020169284A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C27
CURRENT APPLICATION NUMBER: US/09/978,697
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29

; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 33.0%; Score 468.5; DB 9; Length 453;
Best Local Similarity 38.1%; Pred. No. 4.5e-34;
Matches 91; Conservative 40; Mismatches 97; Indels 11; Gaps 6;

QY 30 YNORIVGENAVPGSWPQVSLQDNTGFHFCGSLSPNWNVTAHC---QVTPGRHFVY 86
DB 213 YSSRIVGNNLSLQSPWQASLQFQ-GYHLCCGGSVITPLWITAAHCYVDLYLPKSWTIQ 271
QY 87 LGYDRSSNAEPQVLSIARATHPNWNNANTMNDLTLKLASPARYTAQVSPVCLASTN 146
DB 272 VGLVSLLDNAPSHL--VEKIVYHSYKPKRLGNDALMKLAGPLTFNEMIQPVCLPNSF 329
QY 147 EALPSGLTCVTTGWGRISGVNVTARLQOVVLPVTVNQC--ROYWGARTIDAMTCAG- 203
DB 330 ENFPDGKVCWTSGWATEDGGDASPV-LNHAAPVPLISKNICNRDHYGGIISPSMLCAGY 388
QY 204 -GSGASSCCDGGPLVCCQKNTWLVIGVSMCTKNCNTQAPAMYTRVSKFTWVNOVM 261
DB 389 LTGGVDSQCGDGGPLVCCQKNTWLVIGVSMCTKNCNTQAPAMYTRVSKFTWVNOVM 447

RESULT 13
US-09-978-192A-69
; Sequence 69, Application US/09978192A
; Patent No. US2002017553A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J
; APPLICANT: Kijavir, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630PIC9
; CURRENT APPLICATION NUMBER: US/09/978, 192A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26

[illegible]

```

Query Match      33.0%; Score 468.5; DB 12; Length 453;
Best Local Similarity 38.1%; Pred. No. 4.5e-34;
Matches 91; Conservative 40; Mismatches 97; Indels 11; Gaps 6;

QY 30 YNORIVNGENAVPGSPWQVYSLQDNTGFHFCGGSLSIPNWVVAHHC---QVTPGRHEV 86
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 213 YSSRIYVGGNNLSLLSQWPQASLOFQ-CYHLCGGSVIPLWITIAAHCCVYDLYLPKSWTIQ 271
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 87 LGSYDRSSNAEPQVLSIARAIATHPNNANTMMNDLTLLKLASPARVTAQVSPVCLASTN 146

```

[illegible]

Search completed: December 20, 2002, 15:21:07
Job time : 9.34783 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run On: December 20, 2002, 15:14:53 ; Search time 14.9565 Seconds
(without alignments)
519.349 Million cell updates/sec

Title: US-09-856-319B-4

Perfect score: 1418

Sequence: 1 MLLSLTLTLVLLSSWGCG.....AMTRYSKFSTWVNOYMAYN 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	755.5	53.3	231	2	US-09-027-337-6
2	755.5	53.3	231	4	US-09-644-600-6
3	750.5	52.9	230	4	US-08-944-483-62
4	737	52.0	229	2	US-08-557-146-13
5	737	52.0	229	2	US-09-154-344-13
6	724.5	51.1	228	1	US-08-278-091-10
7	724.5	51.1	228	1	US-08-483-859-10
8	724.5	51.1	228	1	US-08-472-173-10
9	724.5	51.1	228	2	US-08-487-167-10
10	724.5	51.1	228	2	US-08-482-816-10
11	724.5	51.1	228	2	US-08-286-149-10
12	724.5	51.1	228	2	US-08-801-499-10
13	724.5	51.1	228	2	US-08-615-271-10
14	724.5	51.1	228	3	US-09-074-660-10
15	724.5	51.1	228	3	US-09-074-659-10
16	724.5	51.1	228	3	US-09-106-468-10
17	724.5	51.1	228	4	US-09-106-466A-10
18	724.5	51.1	228	4	US-09-106-467-10
19	489	34.5	814	1	US-08-750-711-1
20	489	34.3	241	4	US-08-944-483-59
21	485.5	34.2	812	1	US-08-248-629A-1
22	485.5	34.2	812	1	US-08-451-932-1
23	485.5	34.2	812	1	US-08-452-260-1
24	485.5	34.2	812	1	US-08-326-785-1
25	485.5	34.2	812	2	US-08-612-788-1
26	485.5	34.2	812	2	US-08-605-598B-1
27	485.5	34.2	812	2	US-08-429-743-1

28 485.5 34.2 812 2 US-08-866-735-1 Sequence 1, Appli
29 485.5 34.2 812 3 US-09-066-028-1 Sequence 1, Appli
30 485.5 34.2 812 5 PCT-US95-05107-1 Sequence 1, Appli
31 485 34.2 791 1 US-08-643-219-1 Sequence 1, Appli
32 485 34.2 791 3 US-08-851-350-1 Patent No. 5200340
33 484 34.1 546 6 5200340-6 Sequence 54, Appl
34 484 34.1 790 2 US-08-469-486-54 Sequence 54, Appl
35 484 34.1 790 2 US-08-469-658-54 Sequence 1, Appli
36 484 34.1 791 2 US-09-131-995-1 Sequence 1, Appli
37 484 34.1 791 2 US-08-832-087B-1 Sequence 1, Appli
38 484 34.1 791 4 US-09-132-154-1 Sequence 1, Appli
39 484 34.1 810 1 US-07-854-603-2 Sequence 2, Appli
40 484 34.1 810 1 US-08-147-000B-29 Sequence 29, Appli
41 484 34.1 810 4 US-09-086-514-1 Sequence 1, Appli
42 484 34.1 810 6 5200340-8 Patent No. 5200340
43 483.5 34.1 268 1 US-08-568-031-2 Sequence 2, Appli
44 483.5 34.1 268 2 US-08-966-319-2 Sequence 2, Appli
45 483.5 34.1 268 3 US-09-153-304-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-027-337-6

; Sequence 6, Application US/09027337B

; Patent No. 5972616

; GENERAL INFORMATION:

; APPLICANT: O'Brien, Timothy J.

; APPLICANT: Tanimoto, Hirotooshi

; TITLE OF INVENTION: TAGD-15: An Extracellular Serine Protease Overexpressed in Breast and Ovarian Carcinomas

; FILE REFERENCE: D6064

; CURRENT APPLICATION NUMBER: US/09/027.337B

; CURRENT FILING DATE: 1998-02-20

; NUMBER OF SEQ ID NOS: 13

; SEQ ID NO 6

; LENGTH: 231

; TYPE: PRT

; ORGANISM: Unknown

; FEATURE:

; OTHER INFORMATION: Serine protease catalytic domain of chymotrypsin (Chymb)

; OTHER INFORMATION: homologous to similar domain in TAGD-15

US-09-027-337-6

Query Match 53.3%; Score 755.5; DB 2; Length 231;
Best Local Similarity 58.2%; Pred. No. 5.7e-64;
Matches 135; Conservative 35; Mismatches 61; Indels 1; Gaps 1;

Qy 33 RIVGCVNAVPGSWPQVSLQDNTGTFHFCGGLISPNWVTAHCVTPGRHVVLEGEYDR 92

Db 1 RIVGCVNAVPGSWPQVSLQDNTGTFHFCGGLISPNWVTAHCVTPGRHVVLEGEYDR 92

Qy 93 SSNAEPVQVLSSTARITAPNNMNTNNDTLKLLKASPARVTAQVSPVCLASTNEALPSG 152

Db 60 GSDEENIQVLAKVFKPKFSILTNNIDITLLKATAPRSQTVSAVCLPSADDDFFAG 119

Qy 153 LTCVVTGNGRISGVGNVTPARLQVVLPTVTVNQCRQYWGARTIDAMICAGSGASSCOG 212

Db 120 TLCATTGNGKTKYNANKTPDKLQQAALLSNAECKSKWGRITDVMICAGASGVSSCMG 179

Qy 213 DSGGLVLCOKGNTVLLIGVSGTKNCNQAPAMTRYSKFSTWVNOYMAYN 264

Db 180 DSGGLVLCOKGAMTVLVIGVSGWSDTCTSSPGVYARVTKLIPWVOKILAN 231

RESULT 2

US-09-644-600-6

; Sequence 6, Application US/09644600

; Patent No. 6451500

; GENERAL INFORMATION:

; APPLICANT: O'Brien, Timothy J.

; APPLICANT: Tanimoto, Hirotooshi

1; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease
2; FILE REFERENCE: Overexpressed in Carcinomas
3; CURRENT APPLICATION NUMBER: US/09/644,600
4; CURRENT FILING DATE: 2000-08-23
5; PRIOR APPLICATION NUMBER: 09/421,213
6; PRIOR FILING DATE: 1999-10-20
7; PRIOR APPLICATION NUMBER: 09/027,337
8; PRIOR FILING DATE: 1998-02-20
9; NUMBER OF SEQ ID NOS: 98
10; SEQ ID NO 6
11; LENGTH: 231
12; TYPE: PRT
13; ORGANISM: Homo sapiens
14; FEATURE:
15; OTHER INFORMATION: Chymotrypsin
16; US-09-644-600-6

Query Match 53.3%; Score 755.5; DB 4; Length 231;
Best Local Similarity 58.2%; Pred. No. 5.7e-64;
Matches 135; Conservative 35; Mismatches 61; Indels 1; Gaps 1;
QY 33 RIVNGENAVPGSPWQVSLQDNTGFHFCGSLISPNWVTAHCVTPGRHFVVLGEYDR 92
Db 1 RIVNGEDAVPGSPWQVSLQDNTGFHFCGSLISEDMVVTAAHCGVRTS-DVVVAGEFDQ 59
QY 93 SSNAEPQVLSIARATHPHNWNANTMNDTLTLKLASPARYTAQVSPVCLASTNEALPSG 152
Db 60 GSDEENIQVLTKAKVFNKFSLLTVNNDITLLKLATPARFSQTVSAVCLPSADDDFPAG 119
QY 153 LTCVTTGWRISGVNTPARLQOVVPLVTVNQCROYGARITDAMICAGSGSSCOG 212
Db 120 TLCATTGWGKTKYNAKTPDKLQQAALPLLSNAECKKSGRRITDVMICAGAGSVSCMG 179
QY 213 DSGPLVLCOKGNTWLVIGVSGTKNCKNIQAPAMYTRVSKFSTWVNOVMAYN 264
Db 180 DSGPLVLCOKGNTWLVIGVSGTCTSSPGVYARVTKLIPWVKILAA 231

RESULT 3
US-08-944-483-62
; Sequence 62, Application US/08944483
; Patent No. 6232456
; GENERAL INFORMATION:
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STEWART, KENT D.
; APPLICANT: STROUPE, STEVEN D.
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
; TITLE OF INVENTION: OF THE PROSTATE
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,483
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:

1; APPLICATION NUMBER:
2; FILING DATE:
3; ATTORNEY/AGENT INFORMATION:
4; NAME: Becker, Cheryl L.
5; REGISTRATION NUMBER: 35,441
6; REFERENCE/DOCKET NUMBER: 6183.US.01
7; TELECOMMUNICATION INFORMATION:
8; TELEPHONE: 847/935-1729
9; TELEFAX: 847/938-2623
10; TELEX:
11; INFORMATION FOR SEQ ID NO: 62:
12; SEQUENCE CHARACTERISTICS:
13; LENGTH: 230 amino acids
14; TYPE: amino acid
15; STRANDEDNESS: single
16; TOPOLOGY: linear
17; MOLECULE TYPE: No. 6232456e
18; US-08-944-483-62
Query Match 52.9%; Score 750.5; DB 4; Length 230;
Best Local Similarity 58.0%; Pred. No. 1.7e-63;
Matches 134; Conservative 35; Mismatches 61; Indels 1; Gaps 1;
QY 34 IVNGENAVPGSPWQVSLQDNTGFHFCGSLISPNWVTAHCVTPGRHFVVLGEYDRS 93
Db 1 IVNGEDAVPGSPWQVSLQDNTGFHFCGSLISEDMVVTAAHCGVRTS-DVVVAGEFDQ 59
QY 94 SSNAEPQVLSIARATHPHNWNANTMNDTLTLKLASPARYTAQVSPVCLASTNEALPSGL 153
Db 60 SDEENIQVLTKAKVFNKFSLLTVNNDITLLKLATPARFSQTVSAVCLPSADDDFPAGT 119
QY 154 TCVTTGWRISGVNTPARLQOVVPLVTVNQCROYGARITDAMICAGSGSSCOGD 213
Db 120 LCATTGWGKTKYNAKTPDKLQQAALPLLSNAECKKSGRRITDVMICAGAGSVSCMGD 179
QY 214 SGGPLVLCOKGNTWLVIGVSGTKNCKNIQAPAMYTRVSKFSTWVNOVMAYN 264
Db 180 SGGPLVLCOKGNTWLVIGVSGTCTSSPGVYARVTKLIPWVKILAA 230

RESULT 4
US-08-557-146-13
; Sequence 13, Application US/08557146
; Patent No. 5834290
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hansson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,146
; FILING DATE: 14-DEC-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113

INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 229 amino acids
TYPE: amino acids
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: polypeptide
US-09-856-319b-4.13

Query Match 52.0%; Score 737; DB 2; Length 229;
Best Local Similarity 57.6%; Pred. No. 3.2e-62;
Matches 133; Conservative 36; Mismatches 60; Indels 2; Gaps 2;
Qy 34 IVNGENAVPGSWPQVSLQDNTGTFHFCGSLISPNWVYTAARHCQVTPGRHFVVLGEYDRS 93
Db 1 IVNGEDAVPGSWPQVSLQDNTGTFHFCGSLISPDWVYTAARHCQVTS-DVVAGEFDQ 59
Qy 94 SNAEPQVLSIARAITHPNWNTMNDLTLKLSAPRYTAQVSPVCLASTNEALPSGL 153
Db 60 SDEENIQVLKIAKVEKNPKFSILTAVNNDITLLKLTAPRFSTVSAVCLPSADDDFPAGT 119
Qy 154 TCVTTGGRISGVNVPARLQOVVPLVTVNQCRQYNGARITDAMICAGSGASSCGD 213
Db 120 LCATTGKTKYNANKTPDKLQQAALPLLSNAECKSKGRITDVMICA-GAGVSSCMGD 178
Qy 214 SGGPLVCQKGNWTLGIVSGTKNCNIQAPAMYTRVSKSTWNOVNAVYN 264
Db 179 SGGPLVCQKGNWTLGIVSGTKNCSTSSPGVYARVTKLIPWQKILAAAN 229

RESULT 5
US-09-154-344-13
Sequence 13, Application US/09154344
Patent No. 5981256
GENERAL INFORMATION:
APPLICANT: Egelrud, Torbjorn
APPLICANT: Hansson, Lennart
TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: White & Case, Patent Department
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/154,344
FILING DATE: 16-SEP-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/557,146
FILING DATE: 14-DEC-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sturner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-181
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 229 amino acids
TYPE: amino acids
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: polypeptide
US-09-154-344-13

Query Match 52.0%; Score 737; DB 2; Length 229;
Best Local Similarity 57.6%; Pred. No. 3.2e-62;
Matches 133; Conservative 36; Mismatches 60; Indels 2; Gaps 2;
Qy 34 IVNGENAVPGSWPQVSLQDNTGTFHFCGSLISPNWVYTAARHCQVTPGRHFVVLGEYDRS 93
Db 1 IVNGEDAVPGSWPQVSLQDNTGTFHFCGSLISPDWVYTAARHCQVTS-DVVAGEFDQ 59
Qy 94 SNAEPQVLSIARAITHPNWNTMNDLTLKLSAPRYTAQVSPVCLASTNEALPSGL 153
Db 60 SDEENIQVLKIAKVEKNPKFSILTAVNNDITLLKLTAPRFSTVSAVCLPSADDDFPAGT 119
Qy 154 TCVTTGGRISGVNVPARLQOVVPLVTVNQCRQYNGARITDAMICAGSGASSCGD 213
Db 120 LCATTGKTKYNANKTPDKLQQAALPLLSNAECKSKGRITDVMICA-GAGVSSCMGD 178
Qy 214 SGGPLVCQKGNWTLGIVSGTKNCNIQAPAMYTRVSKSTWNOVNAVYN 264
Db 179 SGGPLVCQKGNWTLGIVSGTKNCSTSSPGVYARVTKLIPWQKILAAAN 229

RESULT 6
US-08-278-091-10
Sequence 10, Application US/08278091
Patent No. 5506139
GENERAL INFORMATION:
APPLICANT: LOOSMORE, Sheena M
APPLICANT: YANG, Yan-ping
APPLICANT: CHONG, Pele
APPLICANT: OOMEN, Raymond P.
APPLICANT: KLEIN, Michel H.
TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
NUMBER OF SEQUENCES: 23
TITLE OF INVENTION: Reduced Protease Activity
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/278,091
FILING DATE: 21-JUL-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-371
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 228 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-278-091-10

Query Match 51.1%; Score 724.5; DB 1; Length 228;
Best Local Similarity 57.1%; Pred. No. 4.9e-61;
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;
Qy 34 IVNGENAVPGSWPQVSLQDNTGTFHFCGSLISPNWVYTAARHCQVTPGRHFVVLGEYDRS 93

Db 1 IVNGEAVPGSPWQVSLQDKTGFHFCGSLINENWVVTAAHCGVTTSDVVVAGEFDQG 59
QY 94 SNAEPVQVLSIARAIHPNNANTMNDLTLLKLSPARYTAQVSPVCLASTNEALPSGL 153
Db 60 SSKSEKIQKIAKFAKNSKYNLSITINDTLTKLSTAAFSQTSVAVCLPSASDDFAAGT 119
QY 154 TCVTTCGWRISGVNVTTPARLQOVLPVTVNOCROYGARITDAMICAGGSGASSCOGD 213
Db 120 TCVTTCGWLTRYAN-TPDRLOQASLPLLSNTNCKKYWGTKIKDAMICAGAGSVSSCMGD 177
QY 214 SGGPLVCQKGNWTLGIVSGTKNCNIQAPAMYTRVSKFSTWINGVMAYN 264
Db 178 SGGPLVCKKNGAWTLGIVSGWSSSTCSTPGVIARVATLVNVMVQOTLAAN 228

RESULT 7

US-08-483-859-10
; Sequence 10, Application US/08483859
; Patent No. 5656436
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: COMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; TITLE OF INVENTION: Reduced Protease Activity
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,859
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-495 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-483-859-10

Query Match 51.1%; Score 724.5; DB 1; Length 228;
Best Local Similarity 57.1%; Pred. No. 4.9e-61;
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;
QY 34 IVNGENAVPGSPWQVSLQDKTGFHFCGSLISPNWVVTAAHCGVTTSPGRHFVVLGEYDRS 93
Db 1 IVNGEAVPGSPWQVSLQDKTGFHFCGSLINENWVVTAAHCGVTTSDVVVAGEFDQG 59

QY 94 SNAEPVQVLSIARAIHPNNANTMNDLTLLKLSPARYTAQVSPVCLASTNEALPSGL 153
Db 60 SSKSEKIQKIAKFAKNSKYNLSITINDTLTKLSTAAFSQTSVAVCLPSASDDFAAGT 119
QY 154 TCVTTCGWRISGVNVTTPARLQOVLPVTVNOCROYGARITDAMICAGGSGASSCOGD 213
Db 120 TCVTTCGWLTRYAN-TPDRLOQASLPLLSNTNCKKYWGTKIKDAMICAGAGSVSSCMGD 177
QY 214 SGGPLVCQKGNWTLGIVSGTKNCNIQAPAMYTRVSKFSTWINGVMAYN 264
Db 178 SGGPLVCKKNGAWTLGIVSGWSSSTCSTPGVIARVATLVNVMVQOTLAAN 228

RESULT 8

US-08-472-173-10
; Sequence 10, Application US/08472173
; Patent No. 5665353
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: COMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; TITLE OF INVENTION: Reduced Protease Activity
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,173
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-493 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-472-173-10

Query Match 51.1%; Score 724.5; DB 1; Length 228;
Best Local Similarity 57.1%; Pred. No. 4.9e-61;
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;
QY 34 IVNGENAVPGSPWQVSLQDKTGFHFCGSLISPNWVVTAAHCGVTTSPGRHFVVLGEYDRS 93
Db 1 IVNGEAVPGSPWQVSLQDKTGFHFCGSLINENWVVTAAHCGVTTSDVVVAGEFDQG 59
QY 94 SNAEPVQVLSIARAIHPNNANTMNDLTLLKLSPARYTAQVSPVCLASTNEALPSGL 153

Db 60 SSSEKIQLKIAKFKNSKYNLSLTINDITLLKLTAAASFQTSVAVCLPSASDDFAAGT 119
QY 154 TCVTTGWRISGVGNVTPARLQOVVPLVTVNQCRQYWGARTTDAMICAGSGASSCGD 213
Db 120 TCVTTGWRG-LTRYAN-TPDRLQOASPLLSNTNCKKYGWTKIKDAMICAGASGVSSCMGD 177
QY 214 SGGPLVCOKGNTWVLIGVSGTKNCNIQAPAMYTRVSKFSTWVNOVMAYN 264
Db 178 SGGPLVCKKNGAWTLVGVSGSSCTSTPGVYARVTVLNVWVQOTLAAN 228

RESULT 9
US-08-487-167-10
: Sequence 10, Application US/08487167
: Patent No. 5869302
: GENERAL INFORMATION:
: APPLICANT: LOOSMORE, Sheena M.
: APPLICANT: YANG, Yan-Ping
: APPLICANT: CHONG, Pele
: APPLICANT: OOMEN, Raymond P.
: APPLICANT: KLEIN, Michel H.
: TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
: TITLE OF INVENTION: Reduced Protease Activity
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sim & McBurney
: STREET: Suite 701, 330 University Avenue
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5G 1R7
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/487.167
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/296.149
: FILING DATE: 26-AUG-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/278.091
: FILING DATE: 21-JUL-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Stewart, Michael I.
: REGISTRATION NUMBER: 24,973
: REFERENCE/DOCKET NUMBER: 1038-508 MIS:vg
: TELEPHONE: (416) 595-1155
: TELEFAX: (416) 595-1163
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 228 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
US-08-487-167-10

Query Match 51.1%; Score 724.5; DB 2; Length 228;
Best Local Similarity 57.1%; Pred. No. 4.9e-61;
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;
QY 34 IVNGENAVPGSPWQVSLQDNTGFHFCGSLISPNWVYTAACQVTPGRHFVVLGEYDRS 93
Db 1 IVNGEAVPGSPWQVSLQDNTGFHFCGSLINWVYTAACQVTPGRHFVVLGEYDRS 59
QY 94 SNAEPVQVLSIARATHPNWNTNNDITLLKLTAAASFQTSVAVCLPSASDDFAAGT 119
Db 60 SSSEKIQLKIAKFKNSKYNLSLTINDITLLKLTAAASFQTSVAVCLPSASDDFAAGT 153
QY 154 TCVTTGWRISGVGNVTPARLQOVVPLVTVNQCRQYWGARTTDAMICAGSGASSCGD 213
Db 60 SSSEKIQLKIAKFKNSKYNLSLTINDITLLKLTAAASFQTSVAVCLPSASDDFAAGT 119

QY 154 TCVTTGWRISGVGNVTPARLQOVVPLVTVNQCRQYWGARTTDAMICAGSGASSCGD 213
Db 120 TCVTTGWRG-LTRYAN-TPDRLQOASPLLSNTNCKKYGWTKIKDAMICAGASGVSSCMGD 177
QY 214 SGGPLVCOKGNTWVLIGVSGTKNCNIQAPAMYTRVSKFSTWVNOVMAYN 264
Db 178 SGGPLVCKKNGAWTLVGVSGSSCTSTPGVYARVTVLNVWVQOTLAAN 228
RESULT 10
US-08-482-816-10
: Sequence 10, Application US/08482816
: Patent No. 5935573
: GENERAL INFORMATION:
: APPLICANT: LOOSMORE, Sheena M
: APPLICANT: YANG, Yan-Ping
: APPLICANT: CHONG, Pele
: APPLICANT: OOMEN, Raymond P.
: APPLICANT: KLEIN, Michel H.
: TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with Reduced Protease
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sim & McBurney
: STREET: Suite 701, 330 University Avenue
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5G 1R7
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/482.816
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/296.149
: FILING DATE: 26-AUG-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/278.091
: FILING DATE: 21-JUL-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Stewart, Michael I.
: REGISTRATION NUMBER: 24,973
: REFERENCE/DOCKET NUMBER: 1038-494 MIS:vg
: TELEPHONE: (416) 595-1155
: TELEFAX: (416) 595-1163
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 228 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
US-08-482-816-10

Query Match 51.1%; Score 724.5; DB 2; Length 228;
Best Local Similarity 57.1%; Pred. No. 4.9e-61;
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;
QY 34 IVNGENAVPGSPWQVSLQDNTGFHFCGSLISPNWVYTAACQVTPGRHFVVLGEYDRS 93
Db 1 IVNGEAVPGSPWQVSLQDNTGFHFCGSLINWVYTAACQVTPGRHFVVLGEYDRS 59
QY 94 SNAEPVQVLSIARATHPNWNTNNDITLLKLTAAASFQTSVAVCLPSASDDFAAGT 119
Db 60 SSSEKIQLKIAKFKNSKYNLSLTINDITLLKLTAAASFQTSVAVCLPSASDDFAAGT 119
QY 154 TCVTTGWRISGVGNVTPARLQOVVPLVTVNQCRQYWGARTTDAMICAGSGASSCGD 213
Db 60 SSSEKIQLKIAKFKNSKYNLSLTINDITLLKLTAAASFQTSVAVCLPSASDDFAAGT 119

Db 120 TCVTGNG-LTRYAN-TPDLQASLPLSNTNCKYWGTHKIDAMICAGASGVSSCMGD 177
QY 214 SGGPLVCQKGNWTLVIGISWGTKNQCIQAPMYTRYSKFSTWNOVMAYN 264
|||||:| | :|:|||||: | :|:| | : | :|:| | : | :|:| | :
Db 178 SGGPLVCCKNGAWTLVIGISWGSSTCSTSPGVYARVTALVNVVQOTLAAN 228

RESULT 11

US-08-296-149-10
; Sequence 10, Application US/08296149
; Patent No. 5939297
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; TITLE OF INVENTION: Reduced Protease Activity
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,149
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-390
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-296-149-10

Query Match 51.1%; Score 724.5; DB 2; Length 228;
Best Local Similarity 57.1%; Pred. No. 4.9e-61;
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;

QY 34 IVGENAVPGSPWQVSLQDNTGFHFCGSLISPNWVYTAHQCQVTPGRHFVVLGEYDRS 93
Db 1 IVNGEAVPGSPWQVSLQDNTGFHFCGSLINENWVYTAHCGVTTSDVVVAGEFDQ 59
|||||:| | :|:|||||: | :|:| | : | :|:| | : | :|:| | :
QY 94 SNAEPVQLSIARATHPNWNTMNDLTLKILASPARTAOVSPVCLASTNEALPSGL 153
Db 60 SSSEKIQKLIKIAKVFKNKSNLSLTINNDITLLKLSAASFQTVSAVCLPSADDFAAGT 119
|||||:| | :|:|||||: | :|:| | : | :|:| | : | :|:| | :
QY 154 TCVTGNGRISGVNTPARLQVPLPVTNOCROYWGARTIDAMICAGSGASSCOGD 213
Db 120 TCVTGNG-LTRYAN-TPDLQASLPLSNTNCKYWGTHKIDAMICAGASGVSSCMGD 177

QY 214 SGGPLVCQKGNWTLVIGISWGTKNQCIQAPMYTRYSKFSTWNOVMAYN 264

Db 178 SGGPLVCCKNGAWTLVIGISWGSSTCSTSPGVYARVTALVNVVQOTLAAN 228

RESULT 12

US-08-801-499-10
; Sequence 10, Application US/08801499
; Patent No. 5962430
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with Reduced Protease A
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,499
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/482,816
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-671 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-801-499-10

Query Match 51.1%; Score 724.5; DB 2; Length 228;
Best Local Similarity 57.1%; Pred. No. 4.9e-61;
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;

QY 34 IVGENAVPGSPWQVSLQDNTGFHFCGSLISPNWVYTAHQCQVTPGRHFVVLGEYDRS 93
Db 1 IVNGEAVPGSPWQVSLQDNTGFHFCGSLINENWVYTAHCGVTTSDVVVAGEFDQ 59
|||||:| | :|:|||||: | :|:| | : | :|:| | : | :|:| | :
QY 94 SNAEPVQLSIARATHPNWNTMNDLTLKILASPARTAOVSPVCLASTNEALPSGL 153
Db 60 SSSEKIQKLIKIAKVFKNKSNLSLTINNDITLLKLSAASFQTVSAVCLPSADDFAAGT 119
|||||:| | :|:|||||: | :|:| | : | :|:| | : | :|:| | :
QY 154 TCVTGNGRISGVNTPARLQVPLPVTNOCROYWGARTIDAMICAGSGASSCOGD 213
Db 120 TCVTGNG-LTRYAN-TPDLQASLPLSNTNCKYWGTHKIDAMICAGASGVSSCMGD 177

QY 214 SGGPLVCQKGNWTLVIGISWGTKNQCIQAPMYTRYSKFSTWNOVMAYN 264

Db 178 SGGPLVCCKNGAWTLVIGISWGSSTCSTSPGVYARVTALVNVVQOTLAAN 228

RESULT 13

US-08-615-271-10
; Sequence 10, Application US/08615271
; Patent No. 5981503
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
; TITLE OF INVENTION: PROTEASE ACTIVITY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/615,271
; FILING DATE: 20-JUN-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-580
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-615-271-10

Query Match 51.1%; Score 724.5; DB 2; Length 228;
Best Local Similarity 57.1%; Pred. No. 4.9e-61;
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;
Qy 34 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAHCVTPGRHFVVLGEYDRS 93
Db 1 IVNGEEAVPGSWPQVSLQDNTGFHFCGSLINENWVYTAHCVTTS-DVVVAGEFDQ 59
Qy 94 SNAEPVQLSTARATHPNWNTNNDLTLKLKSPARYTAQVSPVCLASTNEALPSGL 153
Db 60 SSSEKIQKIAKFKNSKYNLTINNDITLLKLTAAFSQTSVAVCLPSASDDFAAGT 119
Qy 154 TCVTTGWRIGSGVNTVPARLQVVLPLVTYNOCROYGARTIDAMICAGSGSSCCGD 213
Db 120 TCVTTGWR-LTRYAN-TDRLQQAASLPLSNTNCKYKKGTKIKDAMICAGASGVSSCMGD 177
Qy 214 SGGPLVCKKNGAWTLVIGVSGSKTNCNTQAPAMYTRVSKFSTWVQVYLAAN 264
Db 178 SGGPLVCKKNGAWTLVIGVSGSSCTSTPGVYARVYALVNWVQVTLAAN 228
RESULT 14
US-09-074-660-10
; Sequence 10, Application US/09074660
; Patent No. 6020183
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.

; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; TITLE OF INVENTION: Reduced Protease Activity
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/09/074,660
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,167
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-731 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-074-660-10
Query Match 51.1%; Score 724.5; DB 3; Length 228;
Best Local Similarity 57.1%; Pred. No. 4.9e-61;
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;
Qy 34 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAHCVTPGRHFVVLGEYDRS 93
Db 1 IVNGEEAVPGSWPQVSLQDNTGFHFCGSLINENWVYTAHCVTTS-DVVVAGEFDQ 59
Qy 94 SNAEPVQLSTARATHPNWNTNNDLTLKLKSPARYTAQVSPVCLASTNEALPSGL 153
Db 60 SSSEKIQKIAKFKNSKYNLTINNDITLLKLTAAFSQTSVAVCLPSASDDFAAGT 119
Qy 154 TCVTTGWRIGSGVNTVPARLQVVLPLVTYNOCROYGARTIDAMICAGSGSSCCGD 213
Db 120 TCVTTGWR-LTRYAN-TDRLQQAASLPLSNTNCKYKKGTKIKDAMICAGASGVSSCMGD 177
Qy 214 SGGPLVCKKNGAWTLVIGVSGSKTNCNTQAPAMYTRVSKFSTWVQVYLAAN 264
Db 178 SGGPLVCKKNGAWTLVIGVSGSSCTSTPGVYARVYALVNWVQVTLAAN 228
RESULT 15
US-09-074-659-10
; Sequence 10, Application US/09074659
; Patent No. 6025342
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele

Search completed: December 20, 2002, 15:20:37
Job time : 15.9565 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:14:08 ; Search time 32.6957 Seconds
(without alignments)
1663.721 Million cell updates/sec

Title: US-09-856-319b-4
Perfect score: 1418
Sequence: 1 MLLSLTSLVLLGSSWGCG.....AMVTRVSKFSTWINGVMAYN 264

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriapi.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1418	100.0	264	11 Q9ER05	Q9er05 mus musculus
2	1414	99.7	264	11 Q9D7P8	Q9d7p8 mus musculus
3	1408	99.3	264	11 Q9D960	Q9d960 mus musculus
4	1365	96.3	264	11 Q9EQ28	Q9eq28 rattus norv
5	894.5	63.1	261	13 Q9W7Q4	Q9w7q4 paralichthy
6	805.5	56.8	260	13 Q9W7Q3	Q9w7q3 paralichthy
7	797.5	56.2	263	11 Q9DC86	Q9dc86 mus musculus
8	795.5	56.1	263	11 Q9CR35	Q9cr35 mus musculus
9	789.5	55.7	263	11 Q9D8X8	Q9d8x8 mus musculus
10	758.5	53.5	263	13 Q9PWQ6	Q9pwq6 gadus morhu
11	674	47.5	164	11 Q9DC82	Q9dc82 mus musculus
12	552.5	39.0	269	11 Q9CQ52	Q9cq52 mus musculus
13	544.5	38.4	269	11 Q9D7T9	Q9d7t9 mus musculus
14	533.5	37.6	266	13 Q9Z077	Q9z077 gadus morhu
15	524.5	37.0	260	13 Q9W7P9	Q9w7p9 paralichthy
16	510	36.0	267	5 Q9BK47	Q9bk47 luidia foli

17	508.5	35.9	269	6 Q95KW7	Q95kw7 bos taurus
18	503	35.5	249	13 Q9W7Q1	Q9w7q1 paralichthy
19	499	35.2	1524	13 Q91674	Q91674 xenopus lae
20	498	35.1	266	13 Q9W7Q0	Q9w7q0 paralichthy
21	494	34.8	270	13 Q91039	Q91039 gadus morhu
22	488.5	34.4	812	11 Q91WJ5	Q91wj5 mus musculus
23	485	34.2	810	4 Q15146	Q15146 homo sapien
24	484.5	34.2	270	4 Q96QL8	Q96ql8 homo sapien
25	483	34.1	269	4 Q96GV5	Q96gv5 homo sapien
26	473	33.4	277	5 Q96899	Q96899 scolopendra
27	470.5	33.2	257	6 Q19023	Q19023 macaca mula
28	468	33.0	334	6 Q46507	Q46507 papio hamad
29	467	32.9	268	13 Q9W7Q2	Q9w7q2 paralichthy
30	466.5	32.9	339	11 Q99L44	Q99l44 mus musculus
31	466.5	32.9	812	11 Q9R0W3	Q9r0w3 rattus norv
32	465.5	32.8	273	11 Q921N4	Q921n4 mus musculus
33	456	32.2	806	6 Q18783	Q18783 macropus eu
34	455.5	32.1	264	13 Q8QGf6	Q8qgf6 xenopus lae
35	455	32.0	453	11 Q8VDE0	Q8vde0 mus musculus
36	454	32.0	454	6 Q46506	Q46506 papio hamad
37	452.5	31.9	247	13 Q42608	Q42608 petromyzon
38	451.5	31.8	247	13 Q42158	Q42158 petromyzon
39	450	31.7	266	11 Q91X79	Q91x79 mus musculus
40	449	31.7	266	6 Q46644	Q46644 macaca fasc
41	446.5	31.5	246	11 Q921R9	Q921r9 mus musculus
42	446.5	31.5	331	11 Q8R1A6	Q8rla6 mus musculus
43	445.5	31.4	246	11 Q9QUK9	Q9quk9 mus musculus
44	443.5	31.3	244	13 Q8OGW3	Q8ogw3 anguilla ja
45	442.5	31.2	490	11 Q920K3	Q920k3 rattus norv

ALIGNMENTS

RESULT 1

ID	Q9ER05	PRELIMINARY;	PRT;	264 AA.
AC	Q9ER05;			
DT	01-MAR-2001 (Tremblrel. 16, Created)			
DT	01-MAR-2001 (Tremblrel. 16, Last sequence update)			
DT	01-JUN-2002 (Tremblrel. 21, Last annotation update)			
DE	Chymopasins (Chymotrypsin A CTRA-1).			
GN	CTRL OR CTRL.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Mitsui S., Yamaguchi N.;			
RT	"Molecular cloning of mouse chymopasins."			
RL	Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=129S6/SVEV7AC; TISSUE=SPLEEN;			
RA	Bjoernslett M.;			
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.			
CC	- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.			
DR	EMBL; AB016228; BAB2075.1; -			
DR	EMBL; AF236365; AAL11034.1; -			
DR	HSSP; P00766; 4CHA.			
DR	MEROPS; S01.256; -			
DR	MGD; MGI:88558; Ctrl.			
DR	InterPro; IPR001314; Chymotrypsin.			
DR	InterPro; IPR001254; Ser_protease_Try.			
DR	Pfam; PF00089; trypsin; 1.			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	SMART; SM00020; Tryp_SPC; 1.			
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.			
DR	PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.			
DR	PROSITE; PS00135; TRYPSIN_SER; 1.			
KW	Hydrolase; Serine protease.			

```

SQ SEQUENCE 264 AA; 28135 MW; 1D979709A07056C2 CRC64;
Query Match 100.0%; Score 1418; DB 11; Length 264;
Best Local Similarity 100.0%; Pred. No. 1.4e-118;
Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLLSLTSLVLLGSSWCGVPAITPALSYNQRIYNGENAVPGSWPQVSLQDNTGFHFC 60
Db 1 MLLSLTSLVLLGSSWCGVPAITPALSYNQRIYNGENAVPGSWPQVSLQDNTGFHFC 60
QY 61 GGSLSIPNWWVTAACHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARATHPNWNTMNN 120
Db 61 GGSLSIPNWWVTAACHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARATHPNWNTMNN 120
QY 121 DLTLLKLPARYTAQVSPVCLASTNEALPSGLTCTVTTGWGRISGVGNVTPARLQOVVLP 180
Db 121 DLTLLKLPARYTAQVSPVCLASTNEALPSGLTCTVTTGWGRISGVGNVTPARLQOVVLP 180
QY 181 LVTVNQCROYGARTIDAMICAGSGASSCGSDGSGPLVCQKNTWVLIGIVSWGTKNCN 240
Db 181 LVTVNQCROYGARTIDAMICAGSGASSCGSDGSGPLVCQKNTWVLIGIVSWGTKNCN 240
QY 241 IQAPAMYTRVSKFSTWVNOVMAYN 264
Db 241 IQAPAMYTRVSKFSTWVNOVMAYN 264
RESULT 2
Q9D7P8 PRELIMINARY; PRT; 264 AA.
AC Q9D7P8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 1810004D15Rik protein.
GN CTRL OR 1810004D15Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sakai H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AK009019; BAB26029.1; -.
DR HSSP; P00766; 4CHA.
DR MEROPS; S01.256; -.
DR MGD; MGI:88558; Ctrl.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
```

```

DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRY_P_SPC; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28167 MW; 1D979469A07056C2 CRC64;
Query Match 99.7%; Score 1414; DB 11; Length 264;
Best Local Similarity 99.6%; Pred. No. 3.1e-118;
Matches 263; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MLLSLTSLVLLGSSWCGVPAITPALSYNQRIYNGENAVPGSWPQVSLQDNTGFHFC 60
Db 1 MLLSLTSLVLLGSSWCGVPAITPALSYNQRIYNGENAVPGSWPQVSLQDNTGFHFC 60
QY 61 GGSLSIPNWWVTAACHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARATHPNWNTMNN 120
Db 61 GGSLSIPNWWVTAACHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARATHPNWNTMNN 120
QY 121 DLTLLKLPARYTAQVSPVCLASTNEALPSGLTCTVTTGWGRISGVGNVTPARLQOVVLP 180
Db 121 DLTLLKLPARYTAQVSPVCLASTNEALPSGLTCTVTTGWGRISGVGNVTPARLQOVVLP 180
QY 181 LVTVNQCROYGARTIDAMICAGSGASSCGSDGSGPLVCQKNTWVLIGIVSWGTKNCN 240
Db 181 LVTVNQCROYGARTIDAMICAGSGASSCGSDGSGPLVCQKNTWVLIGIVSWGTKNCN 240
QY 241 IQAPAMYTRVSKFSTWVNOVMAYN 264
Db 241 IQAPAMYTRVSKFSTWVNOVMAYN 264
RESULT 3
Q9D960 PRELIMINARY; PRT; 264 AA.
ID Q9D960;
AC Q9D960;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 1810004D15Rik protein.
GN CTRL OR 1810004D15Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sakai H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AK007333; BAB24967.1; -.
```

```
DR HSSP; P00766; 4CHA.
DR MEROPS; S01.256; -.
DR MGD; MGI:88558; Ctrl.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28151 MW; 1D979719E07C16DE CRC64;

Query Match 99.3%; Score 1408; DB 11; Length 264;
Best Local Similarity 99.6%; Pred. No. 1.1e-117;
Matches 263; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLLSLTSLVLLGSSWCGVPAITPALSYNQRIYNGENAVPGSWPQVLSQDNTGFHFC 60
Db 1 MLLSLTSLVLLGSSWCGVPAITPALSYNQRIYNGENAVPGSWPQVLSQDNTGFHFC 60

Qy 61 GGSLSIPNNVYTAHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARATHPHNNANTMNN 120
Db 61 GGSLSIPNNVYTAHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARATHPHNNANTMNN 120

Qy 121 DLTLLKASPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQOVVLP 180
Db 121 DLTLLKASPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQOVVLP 180

Qy 181 LVTNQCQRYWGARTIDAMICAGSGASSCGDGGPLVCQKGNVTWVLIGIVSWGTKNCN 240
Db 181 LVTNQCQRYWGARTIDAMICAGSGASSCGDGGPLVCQKGNVTWVLIGIVSWGTKNCN 240

Qy 241 IOAPAMYTRVSKFTWVNOVMAYN 264
Db 241 IOAPAMYTRVSKFTWVNOVMAYN 264

RESULT 4
Q9EQZ8 PRELIMINARY; PRT; 264 AA.
ID Q9EQZ8
AC Q9EQZ8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chymopasin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=RAT PANCREAS;
RA Sogame Y., Mitsui S., Kataoka K., Kashima K., Kato M., Sakagami J.,
RA Yamaguchi N.;
RT "Molecular cloning of rat chymopasin.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB020757; BAB20287.1; -.
DR HSSP; P00766; 4CHA.
DR MEROPS; S01.256; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28116 MW; F9ED5D210FD3500E CRC64;
```

```
Query Match 96.3%; Score 1365; DB 11; Length 264;
Best Local Similarity 93.6%; Pred. No. 7.2e-114;
Matches 247; Conservative 16; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLLSLTSLVLLGSSWCGVPAITPALSYNQRIYNGENAVPGSWPQVLSQDNTGFHFC 60
Db 1 MLLSLTSLVLLGSSWCGVPAITPALSYNQRIYNGENAVPGSWPQVLSQDNTGFHFC 60

Qy 61 GGSLSIPNNVYTAHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARATHPHNNANTMNN 120
Db 61 GGSLSIPNNVYTAHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARATHPHNNANTMNN 120

Qy 121 DLTLLKASPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQOVVLP 180
Db 121 DLTLLKASPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQOVVLP 180

Qy 181 LVTNQCQRYWGARTIDAMICAGSGASSCGDGGPLVCQKGNVTWVLIGIVSWGTKNCN 240
Db 181 LVTNQCQRYWGARTIDAMICAGSGASSCGDGGPLVCQKGNVTWVLIGIVSWGTKNCN 240

Qy 241 IOAPAMYTRVSKFTWVNOVMAYN 264
Db 241 IOAPAMYTRVSKFTWVNOVMAYN 264

RESULT 5
Q9W7Q4 PRELIMINARY; PRT; 261 AA.
ID Q9W7Q4
AC Q9W7Q4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chymotrypsinogen 1.
OS Paralicthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidae; Paralicthyidae; Paralicthys.
OX NCBI_TaxID=8255;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Suzuki T., Srivastava A.S., Kurokawa T.;
RT "Japanese flounder mRNA for chymotrypsinogen 1.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB029753; BAA82365.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.256; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 261 AA; 28184 MW; D7090A9D65395B7D CRC64;

Query Match 63.1%; Score 894.5; DB 13; Length 261;
Best Local Similarity 63.0%; Pred. No. 6.6e-72;
Matches 167; Conservative 33; Mismatches 60; Indels 5; Gaps 3;
```

```
Qy 1 MLLSLTSLVLLGSSWCGVPAITPALSYNQRIYNGENAVPGSWPQVLSQDNTGFHFC 60
Db 1 MLLFISCFALV--ASALGCGVPSIKPQVSYGKYNGETAVSGSWPQVLSQDGRGFHFC 58

Qy 61 GGSLSIPNNVYTAHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARATHPHNNANTMNN 120
Db 59 GGSLSIPYVWVTAHCTVSPRHRHVLGSDHSDQYNEPIQVMSIARATHPHYNSQNFNN 118
```



```

Db 180 IYSEAKCKESGSKITDVMICAGASGVSSCMGDSGGPLVCQKDGWTLGIVSGSGFCS 239
Qy 241 IQAPAMTRVSKFSTWVNOVMAYN 264
Db 240 TSTPAVYARVATLMPWQOEILEAN 263

RESULT 8
Q9CR35 Q9CR35 PRELIMINARY; PRT; 263 AA.
AC Q9CR35;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE 2200008D09Rik protein.
GN 2200008D09Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=STOMACH, SPLEEN, AND PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AK008927; BAB25971.1; -
DR EMBL; AK003079; BAB22553.1; -
DR EMBL; AK007765; BAB25241.1; -
DR EMBL; AK007815; BAB25280.1; -
DR EMBL; AK008729; BAB25861.1; -
DR EMBL; AK008888; BAB25954.1; -
DR HSSP; P00766; LGCT.
DR MEROPS; S01.152; -.
DR MGD; MGI:1913723; 2200008D09Rik.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydrolase; Serine protease.
SQ SEQUENCE 263 AA; 27822 MW; 28C4487AF1A36B27 CRC64;

Query Match 56.18; Score 795.5; DB 11; Length 263;
Best Local Similarity 54.98; Pred. No. 4.5e-63;
Matches 145; Conservative 41; Mismatches 77; Indels 1; Gaps 1;

Qy 1 MLLSLTLSSLVLLGSSCGGVPATPALSYNQRIYNGENAVPGSWPQVSLQDNTGFHFC 60

```

```

Db 1 MAFLWLVSCFALVGATFCGVPATPVLTLGLSRIVNGEDATPGSWPQVSLQDNTGFHFC 60
Qy 61 GGSLLSPNWWVYTAHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARATHNPNWNTWNN 120
Db 61 GGSLLSPNWWVYTAHCQVTK-TDVVVAGEFDGSDDEENQVVKIAQVFKPNKFSFTVRN 119
Qy 121 DLTLLKLASPARYTAQVSPCLASTNEALPSGLCTVTTGWRISGVGNVTPARLQOVVLP 180
Db 120 DITLLKLATPAQFSETVSAVCLPTVDDFFPATGTCATGKTKYNAUKTDPKLOOAAALP 179
Qy 181 LVTYNQCRQYWGARTIDAMICAGSGGSSCGDGGPLVCQKNGTWNVLIGVSWGTKNCN 240
Db 180 IVSEAKCKESGSKITDVMICAGASGVSSCMGDSGGPLVCQKDGWTLGIVSGSGFCS 239
Qy 241 IQAPAMTRVSKFSTWVNOVMAYN 264
Db 240 TSTPAVYARVATLMPWQOEILEAN 263

RESULT 9
Q9D8X8 Q9D8X8 PRELIMINARY; PRT; 263 AA.
AC Q9D8X8;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE 2200008D09Rik protein.
GN 2200008D09Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AK007566; BAB25112.1; -
DR HSSP; P00766; LGCT.
DR MEROPS; S01.152; -.
DR MGD; MGI:1913723; 2200008D09Rik.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydrolase; Serine protease.
SQ SEQUENCE 263 AA; 27898 MW; C0638FB8F905A92F CRC64;

```

Query Match 55.7%; Score 789.5; DB 11; Length 263;

Best Local Similarity 54.5%; Pred. No. 1.6e-62;

Matches 144; Conservative 41; Mismatches 78; Indels 1; Gaps 1;

QY 1 MLLSLTSLVLLGSSWGGVPAITPALSYNQRIVNGENAVPGSWPQVSLQDNTGCFHFC 60
 Db 1 MAFLLVSCFALVATFGCGVPAIQVLTGLSRVNGEDAIPGSPWQVSLQDRTGCFHFC 60
 QY 61 GGLISPNWVVAHQCQVTPGRHFVVLGVDYDRSSNAEPQVLSIARAITHPNWNANTMNN 120
 Db 61 GGLISENWWVAHCGVKT-TDVVAGFEFDOGSDEENQVQLKIAQVFNKPFNSFTVRN 119
 QY 121 DITLLKLAPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVNGVTPARLQOVLP 180
 Db 120 DITLLKLATPAQSETVSAVCLPTVDDEPAGTLCATTGKTKYNALKTDPKLOQAALP 179
 QY 181 LVTVNCQYWGARIIDAMICAGGSGASSCGSGGLPVCQKGNVTWVLIGIVSWGTKNCP 240
 Db 180 IVSEAKCKESWGSKITDVMICAGASGVSSCMGDSGGGLPVCQKQDGVWVTLAGIVSWGSGFCS 239
 QY 241 IOAPAMYTRVSKFSTWINOVAYN 264
 Db 240 TSTPAYARVATMPVQVILEAN 263

RESULT 10

Q9PWQ6 PRELIMINARY; PRT; 263 AA.
 AC Q9PWQ6;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2002 (TREMBlrel. 20, Last annotation update)
 DE Chymotrypsin B precursor (EC 3.4.21.1).
 GN CHYB
 OS Gadus morhua (Atlantic cod).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
 OX NCBI_TaxID=8049;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PYLORIC CARCA;
 RX MEDLINE=20464334; PubMed=11011764;
 RA Spilliaert R., Gudmundsdottir A.;
 RT "Molecular Cloning of the Atlantic Cod Chymotrypsinogen B";
 RL Microb. Comp. Genomics 5:41-50(2000).
 CC 1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.
 CC EMBL: AJ242521; CAB43766.1; .
 DR HSSP; P00766; ICHG.
 DR MEROPS; S01.152; .
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser.protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp.Spc; 1.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Signal.
 FT SIGNAL 1 16 POTENTIAL.
 FT CHAIN 17 263 CHYMOTRYPSIN B.
 SQ SEQUENCE 263 AA; 28175 MW; EF61B18A34EE5E7C CRC64;

Query Match 53.5%; Score 758.5; DB 13; Length 263;

Best Local Similarity 53.5%; Pred. No. 9.1e-60;

Matches 137; Conservative 47; Mismatches 65; Indels 7; Gaps 5;

QY 13 LGSSWGGVPAITPALSYNQRIVNGENAVPGSWPQVSLQDNTGPFHFCGSGSLISPNWVY 72
 Db 11 ISAAVCGSPAIQVPTGVARIIVNGEAVPHSWPQVSLQSSQSGFPHFCGSLINENWVY 70
 QY 73 AHCQVTPGRHFVVLGVDYDRS-SNAEPQVLSIARAITHPNWNANTMNDLLKLKASPA 131

Db 71 AHCNVRT-YHRVIVGEHDKSRASDENIQLKPSWFTTHPKWDSRTINNDISLIKASPA 129
 QY 132 RYTAQVSPVCLASTNEALPSGLTCVTTG--RISGVNGVTPARLQOVLPVLTVNOCRQ 189
 Db 130 VLGTVNPPVCLGESSDVFAFGMKCVTSGLTRYNAPG--TPNKLQQAALPLMSNECSQ 187
 QY 190 YWGAR-ITDAMICAGGSGASSCGSGGLPVCQKGNVTWVLIGIVSWGTKNCPAPMYT 248
 Db 188 TWGNMISDVMICAGAAGATSCMGDSGGPLVCQKDNVTLGVISWSSRCSVTTPAVYA 247
 QY 249 RVSKFSTWINOVAYN 264
 Db 248 RVTELGRWVQDILAN 263

RESULT 11

Q9DC82 PRELIMINARY; PRT; 164 AA.
 AC Q9DC82;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Q910001G08RIK protein.
 GN CTRL OR 0910001G08RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=SPLEEN;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustingich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection";
 RL Nature 409:685-690(2001).
 DR EMBL; AK003074; BAB22549.1; .
 DR HSSP; P00763; IDPO.
 DR MEROPS; S01.997; .
 DR MGD; MGI:88558; Ctrl.
 DR InterPro; IPR001254; Ser.protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR SMART; SM00020; Tryp.Spc; 1.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 164 AA; 17707 MW; 83791FD829ABEDD6 CRC64;

Query Match 47.5%; Score 674; DB 11; Length 164;

Best Local Similarity 82.2%; Pred. No. 1.7e-52;

Matches 129; Conservative 5; Mismatches 21; Indels 2; Gaps 2;

QY 110 HPNWNNA-NTMNDLLTLKLKASP-ARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGV 167
 Db 8 HPWGGPHLPKKTRLPLEFFFPPEKTKTTPPPGLASTNEALPSGLTCVTTGWRISGV 67

Db 175 ATCTRSDWNGSLVTNSMYCAGAGDLASCNGDSGGPLACQSPDGSWEVHGYYVFGSSMGC 234
QY 240 NI-QAPAMYTRVSKFSTWINOVM 261
Db 235 NYPKAPSVFTRVSAYIPWINVM 257

Search completed: December 20, 2002, 15:18:54
Job time : 33.6957 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:12:13 ; Search time 40.6957 Seconds
(without alignments)
864.421 Million cell updates/sec

Title: US-09-856-319B-4

Perfect score: 1418

Sequence: 1 MLLSLTSLVLLGSSWGCG.....AMTRYSKFSTWNOVMAYN 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1418	100.0	264	21	Mouse serine prote
2	1262	89.0	264	21	Human serine prote
3	889	62.7	192	21	Human pancreatic c
4	822.5	58.0	263	23	Amino acid sequenc
5	774.5	54.6	245	21	Bovine chymotrypsi
6	755.5	53.3	231	22	Human chymotrypsin
7	537	37.9	269	7	Porcine elastase I
8	537	37.9	269	8	Pig pancreas elast
9	523.5	36.9	253	13	Porcine pancreatic
10	515	36.3	269	7	Sequence of human

11	500	35.3	279	22	ABG20513	Novel human diago
12	498	35.1	269	7	AAP61723	Human elastase II.
13	496.5	35.0	268	17	AAR90683	Human caldecrin co
14	492	34.7	252	7	AAP60058	Sequence of human
15	487.5	34.4	811	12	AAR12933	Plasminogen muten
16	486.5	34.3	268	17	AAR90682	Rat caldecrin cont
17	486.5	34.3	811	12	AAR12939	Plasminogen muten
18	486.5	34.3	811	12	AAR12945	Plasminogen muten
19	485.5	34.2	270	8	AAP70759	Human pancreas ela
20	485.5	34.2	811	12	AAR12943	Plasminogen muten
21	485.5	34.2	812	16	AAR83959	Complete mouse pla
22	485.5	34.2	812	17	AAR07585	Murine plasminogen
23	485.5	34.2	812	20	AAW08686	Murine plasminogen
24	485.5	34.2	812	20	AAW94036	Murine plasminogen
25	485.5	34.2	812	21	AAW00053	Murine plasminogen
26	485.5	34.2	812	21	AAW16490	Murine plasminogen
27	485.5	34.2	812	22	AAW50516	Murine plasminogen
28	485.5	34.2	812	22	AAW67229	Amino acid sequenc
29	485	34.2	791	18	AAW34285	Human plasminogen.
30	485	34.2	810	23	ABB83470	Human plasminogen.
31	485	34.2	810	23	ABB75939	Human plasminogen.
32	485	34.2	1220	22	AAU32858	Novel human secret
33	484.5	34.2	270	23	AAU87691	Human pancreatic t
34	484.5	34.2	273	21	AAW34243	Human pancreatic c
35	484.5	34.2	811	12	AAR12944	Plasminogen muten
36	484	34.1	249	23	ABB81497	Human microplasmin
37	484	34.1	348	23	ABB81498	Human miniplasmin
38	484	34.1	437	19	AAW51457	Human plasminogen
39	484	34.1	546	14	AAR34427	Sequence of tissue
40	484	34.1	790	15	AAR60519	Human 'Glu' plasmi
41	484	34.1	790	22	AAB36562	Mammalian kringl
42	484	34.1	791	21	AAB01887	Human plasminogen.
43	484	34.1	791	21	AAW9589	Human plasminogen.
44	484	34.1	791	21	AAW50867	Human plasminogen.
45	484	34.1	791	22	AAW67223	Amino acid sequenc

ALIGNMENTS

RESULT 1
AAB11711
ID AAB11711 standard; Protein: 264 AA.
XX
AC AAB11711;
XX

DT 23-OCT-2000 (first entry)

XX Mouse serine protease BSSP5 (mBSSP5) SEQ ID NO:4.

DE BSSP5; serine protease; human; hbSSP5; mouse; mBSSP5; brain;
KW diagnostic marker; antibody; transgenic animal; Alzheimer's disease;
KW epilepsy; cancer; inflammation; infertility; pancreatitis;
KW prostatic hypertrophy.

XX Mus sp.

PN WO200031243-A1.

XX 02-JUN-2000.

XX 19-NOV-1999; 99WO-JP06473.

XX 20-NOV-1998; 98JP-0347806.

XX (FUSO) FUSO PHARM IND LTD.

XX Uemura H, Okui A, Kominami K, Yamaguchi N, Mitsui S;

XX WPI; 2000-400058/34.

XX N-PSDB; AAA61734.

XX Serine proteases BSSP5, useful in detecting homologs, mutants and

PT polymorphic variants as markers for diagnosis of e.g. Alzheimer's
PT disease, epilepsy, cancer and inflammation, using blood, urine,
XX pancreas or other tissues
XX
XX Claim 3; Page 55-56; 70pp; Japanese.
XX
CC The invention relates to novel serine proteases designated BSSP5
CC (AAB11710-B11711), and to nucleic acids encoding them (AAA61733-A61734).
CC The invention also relates to vectors and transformants comprising BSSP5
CC nucleic acids; transgenic animals in which the expression level of BSSP5
CC can be varied; and an mBSSP5 knockout mouse. The invention additionally
CC encompasses anti-BSSP5 antibodies and methods of production of such
CC antibodies, methods of BSSP5 detection using the antibodies, and the
CC use of BSSP5 proteins or fragments as diagnostic markers for certain
CC medical conditions, e.g., pancreatitis. A method for detecting
CC pancreatitis comprising measuring BSSP5 concentration in the blood or
CC urine, and a pancreatitis diagnostic agent containing an anti-BSSP5
CC antibody is also disclosed. Nucleotides encoding BSSP5 were initially
CC isolated in a human brain cDNA library using degenerate PCR primers
CC (AAA61744-A61745) based on conserved regions of serine proteases. The
CC BSSP5 serine proteases and nucleotides encoding them are useful in
CC detecting homologues, mutants and polymorphic variants in biological
CC samples (e.g., blood, urine, brain, prostate gland, placenta, testis,
CC pancreas and spleen) as diagnostic markers for conditions such as
CC Alzheimer's disease, epilepsy, cancer, inflammation, infertility,
CC pancreatitis and prostatic hypertrophy. Sequence AAB11710 represents
CC human BSSP5 (hBSSP5), and sequence AAB11711 represents murine BSSP5
CC (mBSSP5).
XX
SQ Sequence 264 AA;

Query Match 100.0%; Score 1418; DB 21; Length 264;
Best Local Similarity 100.0%; Pred. No. 7.9e-112;
Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGSGGCGVPAITPALSYNQRIYNGENAVPGSWPQVSLQDNTGFHFC 60
DB 1 MLLSLTSLVLGSGGCGVPAITPALSYNQRIYNGENAVPGSWPQVSLQDNTGFHFC 60

QY 61 GGSLSIPNNVVTAAHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARAIHPNNANTMNN 120
DB 61 GGSLSIPNNVVTAAHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARAIHPNNANTMNN 120

QY 121 DLTLKLPARTYAQVSPVCLASTNEALPSGLTCVTTGWRISGNGVNTPARLQOVWLP 180
DB 121 DLTLKLPARTYAQVSPVCLASTNEALPSGLTCVTTGWRISGNGVNTPARLQOVWLP 180

QY 181 LVTNOCROYWGARIITDAMICAGSGCGSGGGLVCQKNTWVLIGIVSWGTKNCN 240
DB 181 LVTNOCROYWGARIITDAMICAGSGCGSGGGLVCQKNTWVLIGIVSWGTKNCN 240

QY 241 IQAPAMYTRVSKFSTWVNOVMAYN 264
DB 241 IQAPAMYTRVSKFSTWVNOVMAYN 264

RESULT 2
AAB11710
ID AAB11710 standard; Protein; 264 AA.
XX AAB11710;
AC AAB11710;
XX
DT 23-OCT-2000 (first entry)
DE Human serine protease BSSP5 (hBSSP5) SEQ ID NO:2.
XX BSSP5; serine protease; human; hBSSP5; mouse; mBSSP5; brain;
KW diagnostic marker; antibody; transgenic animal; Alzheimer's disease;
KW epilepsy; cancer; inflammation; infertility; pancreatitis;
XX prostatic hypertrophy.
OS Homo sapiens.
XX

PN WO200031243-A1.
XX
PD 02-JUN-2000.
XX
PF 19-NOV-1999; 99WO-JP06473.
XX
PR 20-NOV-1998; 98JP-0347806.
XX
XX (FUSO) FUSO PHARM IND LTD.
XX
PI Uemura H, Okui A, Kominami K, Yamaguchi N, Mitsui S;
XX
DR WPI; 2000-400058/34.
DR N-PSDB; AAA61733.
XX
XX Serine proteases BSSP5, useful in detecting homologs, mutants and
PT polymorphic variants as markers for diagnosis of e.g. Alzheimer's
PT disease, epilepsy, cancer and inflammation, using blood, urine,
XX pancreas or other tissues
XX
PS Claim 1; Page 51-52; 70pp; Japanese.
XX
CC The invention relates to novel serine proteases designated BSSP5
CC (AAB11710-B11711), and to nucleic acids encoding them (AAA61733-A61734).
CC The invention also relates to vectors and transformants comprising BSSP5
CC nucleic acids; transgenic animals in which the expression level of BSSP5
CC can be varied; and an mBSSP5 knockout mouse. The invention additionally
CC encompasses anti-BSSP5 antibodies and methods of production of such
CC antibodies, methods of BSSP5 detection using the antibodies, and the
CC use of BSSP5 proteins or fragments as diagnostic markers for certain
CC medical conditions, e.g., pancreatitis. A method for detecting
CC pancreatitis comprising measuring BSSP5 concentration in the blood or
CC urine, and a pancreatitis diagnostic agent containing an anti-BSSP5
CC antibody is also disclosed. Nucleotides encoding BSSP5 were initially
CC isolated in a human brain cDNA library using degenerate PCR primers
CC (AAA61744-A61745) based on conserved regions of serine proteases. The
CC BSSP5 serine proteases and nucleotides encoding them are useful in
CC detecting homologues, mutants and polymorphic variants in biological
CC samples (e.g., blood, urine, brain, prostate gland, placenta, testis,
CC pancreas and spleen) as diagnostic markers for conditions such as
CC Alzheimer's disease, epilepsy, cancer, inflammation, infertility,
CC pancreatitis and prostatic hypertrophy. Sequence AAB11710 represents
CC human BSSP5 (hBSSP5), and sequence AAB11711 represents murine BSSP5
CC (mBSSP5).
XX
SQ Sequence 264 AA;

Query Match 89.0%; Score 1262; DB 21; Length 264;
Best Local Similarity 86.0%; Pred. No. 1.2e-98;
Matches 227; Conservative 25; Mismatches 12; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGSGGCGVPAITPALSYNQRIYNGENAVPGSWPQVSLQDNTGFHFC 60
DB 1 MLLSLTSLVLGSGGCGVPAITPALSYNQRIYNGENAVPGSWPQVSLQDNTGFHFC 60

QY 61 GGSLSIPNNVVTAAHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARAIHPNNANTMNN 120
DB 61 GGSLSIPNNVVTAAHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARAIHPNNANTMNN 120

QY 121 DLTLKLPARTYAQVSPVCLASTNEALPSGLTCVTTGWRISGNGVNTPARLQOVWLP 180
DB 121 DLTLKLPARTYAQVSPVCLASTNEALPSGLTCVTTGWRISGNGVNTPARLQOVWLP 180

QY 181 LVTNOCROYWGARIITDAMICAGSGCGSGGGLVCQKNTWVLIGIVSWGTKNCN 240
DB 181 LVTNOCROYWGARIITDAMICAGSGCGSGGGLVCQKNTWVLIGIVSWGTKNCN 240

QY 241 IQAPAMYTRVSKFSTWVNOVMAYN 264
DB 241 VRAPAVYTRVSKFSTWVNOVMAYN 264

RESULT 3

CC over-expressed in ovarian tumours. TADG-15 protein or its fragments of
 CC 9-20 residues that lack TADG-15 protease activity are useful for
 CC vaccinating an individual against TADG-15, having, suspected of having or
 CC at risk of getting cancer. Furthermore, the TADG-15 gene can be used as a
 CC diagnostic or therapeutic target in cancer. The present sequence was used
 CC in a sequence homology alignment with the catalytic domain of TADG-15.

XX Sequence 231 AA;

Query Match 53.3%; Score 755.5; DB 22; Length 231;
 Best Local Similarity 58.2%; Pred. No. 5.8e-56;
 Matches 135; Conservative 35; Mismatches 61; Indels 1; Gaps 1;

QY 33 RIVGENAVPGSWPQVSLQDNTGFHFCGGLISPNVWVTAHCOVTPGRHFVVLGEYDR 92

DB 1 RIVGENAVPGSWPQVSLQDNTGFHFCGGLISPNVWVTAHCOVTPGRHFVVLGEYDR 92

QY 93 SSNAEPVOVLSIARATHPNKNMNTMNDLLKLSAPRYTAQVSPVCLASTNEALPSG 152

DB 60 GSDEENIQVLKIAKVFKNPKFSILTVNNDITLLKATPAREFQTSVAVCLPSADDDPAG 119

QY 153 LTCVTTGWRISGNGVTPARLQVPLVTVNOCROYWGARIITDAMICAGSGASCQG 212

DB 120 TLCATTGKTKYNANKTPDKLQQAALPLLSNAECKKSWGRRITDVMICAGSGVSCMG 179

QY 213 DSGGPLVCQKNTWVLIGISWGTKNICNOIAPAMYTRVSKFSTWINOVMAYN 264

DB 180 DSGGPLVCQKNTWVLIGISWGTKNICNOIAPAMYTRVSKFSTWINOVMAYN 231

RESULT 7.

AAP61724

ID AAP61724 standard; Protein; 269 AA.

XX AAP61724;

XX 23-OCT-1991 (first entry)

XX Porcine elastase II.

XX Colibacillus; yeast.

XX Sus scrofa.

XX JP61192289-A.

XX 26-AUG-1986.

XX 22-FEB-1985; 85JP-0034050.

XX 22-FEB-1985; 85JP-0034050.

XX (KIRI) KIRIN BREWERY KK.

XX WPI; 1986-262895/40.

XX N-PSDB; AAN60919.

XX Biological prepn. of pig elastase II - using DNA with base

PT sequence coding for specified aminoacid sequence.

XX Claim 1; Fig 1; 11pp; Japanese.

XX The porcine elastase product may be efficiently expressed from a
 CC transformed host such as colibacillus or yeast, yielding the product
 CC more efficiently than spleen extraction.
 CC See also J61192288.

XX Sequence 269 AA;

Query Match 37.9%; Score 537; DB 7; Length 269;

Best Local Similarity 43.1%; Pred. No. 2e-37;

Matches 118; Conservative 48; Mismatches 92; Indels 16; Gaps 11;

QY 2 LLLSLTSLVLLGSSWCGVPAITPALSYNORIVNGENAVPGSWPQVSLQ-DNTG--FH 58
 DB 1 MIRALLSTLVAG-ALSCGLPANLPQL--PRVVGEDARPNPWPQVSLQYDSSGQWRH 56
 QY 59 FCGSLSPNNVWVTAHCOVTPGRHFVVLGEYDRSSNAEPVOVLSIARATHPNKNMNTM 118
 DB 57 TCGTSLVDQSVLWVTAHCOVTPGRHFVVLGEYDRSSNAEPVOVLSIARATHPNKNMNTM 116
 QY 119 N--NDLTLLKLSAPRYTAQVSPVCLASTNEALPSGLTCVTTGWRISGNGVTPARLQ 176
 DB 117 SKGNDIALKLSAPRYLTKIOLGCLPAGSILPNVWVTAHCOVTPGRHFVVLGEYDR 175
 QY 177 VVLPLVTVNOCQ--YWGARIITDAMICAGSG--ASSCGDGGPLVCQKGN--TWVLIGIV 232
 DB 176 GQLLVVDYATCSKPCWWSGTVKTNNICAGDGISSCNGDGGPLNCOCANGQVHGIV 235
 QY 233 SWGTK-NCN-TOAPAMYTRVSKFSTWINOVMAYN 264
 DB 236 SFGSSLCGNYVHKPSVETRVSNYIDWINSVIANN 269

RESULT 8

AAP70758

ID AAP70758 standard; Protein; 269 AA.

XX AAP70758;

XX 19-APR-1991 (first entry)

XX Pig pancreas elastase-2.

XX Lipoprotein metabolism; lipid.

XX Sus scrofa.

XX Key Location/Qualifiers

FT Misc-difference 2..16

FT /note= "May be absent"

XX JP62000276-A.

XX 06-JAN-1987.

XX 25-JUN-1985; 85JP-0138494.

XX 25-JUN-1985; 85JP-0138494.

XX (SANY) SANKYO KK.

XX WPI; 1987-040875/06.

XX N-PSDB; AAN71121.

XX Pancreas elastase prodn. used to improve lipoprotein metabolism -

PT comprises isolating RNA coding elastase, synthesising single and

XX double chain cDNA and introducing recombinant into host

XX Disclosure; Page 480; 18pp; Japanese.

XX Sequence encoding pancreas elastase may be used to transform an

XX expression system. The product is useful in rejuvenating the

XX elasticity of the arterial wall, and improving unusual serum lipid

XX levels and lipoprotein metabolism.

XX Sequence 269 AA;

Query Match 37.9%; Score 537; DB 8; Length 269;

Best Local Similarity 43.1%; Pred. No. 2e-37;

Matches 118; Conservative 48; Mismatches 92; Indels 16; Gaps 11;

QY 2 LLLSLTSLVLLGSSWCGVPAITPALSYNORIVNGENAVPGSWPQVSLQ-DNTG--FH 58

DB 1 MIRALLSTLVAG-ALSCGLPANLPQL--PRVVGEDARPNPWPQVSLQYDSSGQWRH 56

QY 233 SWGK-NCN-IOAPMYTRVSKFSTWVNOVMAYN 264
 Db 236 SFGSLGCGNYHKPSVFTVRSNYIDWINSVIANN 269

RESULT 11

ABG20513
 ID ABG20513 standard; Protein: 279 AA.

XX AC ABG20513;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #20504.

XX XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS84700.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

XX PS Claim 20; SEQ ID No 50872; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 279 AA;
 Query Match 35.3%; Score 500; DB 22; Length 279;
 Best Local Similarity 39.9%; Pred. No. 2.7e-34;
 Matches 112; Conservative 45; Mismatches 96; Indels 28; Gaps 11;

QY 2 LLSLTSLVILGSSWGGCVPAITPALSYNQRIVNGENAVPGSWPQVSLQDNTG---FH 58

Db 9 MIRTLLSTLVAG-ALSCGDPYTP---YVTRVVGGEARPNWQVSLQYSSNGKWTYH 64
 QY 59 FCGSLISPNNVWVYTAHC-----QVTPGRHFVVVLGEYDRSSNAEPVQVLSIARITHP 111
 Db 65 TCGGSLIANSWVLTAAHCISSTRTYRVGLGRHNLVVAE-----SGSLAMSSSVKIVVHK 118
 QY 112 NNNANTMN--NDLTLLKLASPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNV 169
 Db 119 DMNFNQISKGNIDIALLLKLANPVSLTKQLACFPAGTILPNYPCTXVTGLGESCRRTNGA 178
 QY 170 TPARLQQVVLPLVTVNOCRQ--YMGARITDAMICAGGSGA-SSCGDSDGGPLVCOQKNT- 225
 Db 179 VPDVLOQGRLLAVDYATCSSSAWGWSSVKTSMICAGGDGVISSCNGDSGGPLNCOASDAL 238
 QY 226 WVLIGIVSWGK-NCN-IOAPAMYTRVSKFSTWVNOVMAYN 264
 Db 239 WQVHGIVSFGSLGCGNYHKPSVFTVRSNYIDWINSVIANN 279

RESULT 12

AAP61723
 ID AAP61723 standard; Protein: 269 AA.

XX AC AAP61723;

XX DT 23-OCT-1991 (first entry)

XX DE Human elastase II.

XX KW Colibacillus; yeast.

XX OS Homo sapiens.

XX PN JP61192288-A.

XX PD 26-AUG-1986.

XX PF 22-FEB-1985; 85JP-0034049.

XX PR 22-FEB-1985; 85JP-0034049.

XX PA (KIRI) KIRIN BREWERY KK.

XX DR WPI; 1986-262894/40.

XX DR N-PSDB; AAN60706.

XX PT Biologically producing human elastase II - using DNA chain with
 biological prodn. capability for human elastase II.

XX PS Claim 1; Fig 1-2; 19pp; Japanese.

XX CC The human elastase product may be efficiently expressed from a
 CC transformed host such as colibacillus or yeast, yielding the product
 CC more efficiently than spleen extraction.
 CC See also J61192289.

XX SQ Sequence 269 AA;

Query Match 35.1%; Score 498; DB 7; Length 269;
 Best Local Similarity 41.2%; Pred. No. 3.9e-34;
 Matches 113; Conservative 46; Mismatches 99; Indels 16; Gaps 10;

QY 2 LLSLTSLVILGSSWGGCVPAITPALSYNQRIVNGENAVPGSWPQVSLQDNTG---FH 58
 Db 1 MIRTLLSTLVAG-ALSCGDPYTP---YVTRVVGGEARPNWQVSLQYSSNGKWTYH 56
 QY 59 FCGSLISPNNVWVYTAHCQVTPGRHFVVVLGEYDRSSNAEPVQVLSIARITHPNNANTM 118
 Db 57 TCGGSLIANSWVLTAAHCISSTRTYRVGLGRHNLVVAESGLAVSVKIVVHKDWSNOI 116
 QY 119 N--NDLTLLKLASPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQ 176

Db	677	VITDKVIPACLPSPNVVVADRTCTCFITNGWETGTGFA--GLLKEAQLPVNIENKVCNRVE	734
QY	191	-WGARIITDAMICAG--CGSASSCQDGGGGLVQCQKNTWYVLIGVSWGKNCNCIQAPMY	247
Db	735	FLNVRVSTELCAGHLAGGTDSCQDGGGLVCFEKDKYILOGVTISWGLGCARPNGVY	794
QY	248	TRVSKFSTWINOVNAYN	264
Db	795	VRVSRFVTWIEGVMRNN	811

Search completed: December 20, 2002, 15:16:41
Job time : 42.6957 secs

RESULT 15
AA12933
ID AA12933 standard; Protein; 811 AA.
XX AC
XX AA12933;
DT 23-SEP-1991 (first entry)
XX Plasminogen mutetin X1 with factor Xa cleavage site.
XX protease; fibrinolysis; blood clotting; thrombosis.
FH Key Location/Qualifiers
FT Cleavage-site 578..581
PN /label= Factor Xa cleavage site
XX WO9109118-A.
XX 27-JUN-1991.
PD 07-DEC-1990; 90WO-G001912.
PF 07-DEC-1989; 89GB-0027722.
PR 07-DEC-1990; 90WG-GB01911.
PA (BRBI-) BRIT BIO-TECHN LTD.
XX Dawson KM, Edwards RM, Forman JM;
PI WPI; 1991-208145/28.
DR N-PSDB; AAQ12542.
XX Activatable fibrinolytic and antithrombin proteins - activated by
PT e.g. factor Xa, thrombin or activated protein C
XX Claim 7; Fig 2 and Fig 4; 73pp; English.
XX This protein is a plasminogen mutant cleavable by Factor Xa.
CC Activation is localised to the thrombus because cleavage to plasmin
CC is by an enzyme of the blood clotting pathway. Compositions
CC comprising the mutant plasminogen are used for treatment or
CC prevention of thrombosis, etc.
CC See AAQ12543-Q12558.
XX Sequence 811 AA;
SQ
Query Match 34.4%; Score 487.5; DB 12; Length 811;
Best Local Similarity 40.9%; Pred. No. 1.1e-32;
Matches 105; Conservative 31; Mismatches 104; Indels 17; Gaps
QY 16 SWGCGVPAITPALSYNORIYNGENAVPGSPHQVSLODNTGFHCFCGSILISPNWVVTAAH 75
Db 564 SFGDCQKQVDFPKICIEIGRVGGCVAPHSPWQVSLRTRFGMHFCGTLLISPEWVLTAAH 623
QY 76 C---QVTPGRHFVVLGEYDRSSNAEP-VQVLTARATHTPNWNANTMNNDLTLKLKASPA 131
Db 624 CLEKSPPSSKYKILGNH-QEVNLEPHGQEIVSRFLFET-----RKDIALLKLSSPA 676
QY 132 RYTAQVSPVCLASTNEALPSGLTCVTTMGRISGVGNVTPARLQQVVLPLVTVNOCROY- 190

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OK protein - protein search, using sw model

Run on: December 20, 2002, 15:16:49 ; Search time 7.30435 Seconds
(without alignments)
530.147 Million cell updates/sec

Title: US-09-856-319B-2_COPY_1_231
Perfect score: 1221
Sequence: 1 MLLSLTSLVLLGSSWGGC.....GDSGGPLVCQKGTWVLIGI 231

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 106657 seqs, 16763532 residues

Total number of hits satisfying chosen parameters: 106657

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	996	81.6	192	10	US-09-925-297-529
2	684.5	56.1	263	10	US-09-888-615-96
3	427.5	35.0	269	10	US-09-925-297-576
4	420.5	34.4	270	10	US-09-923-779-152
5	420.5	34.4	273	10	US-09-925-297-695
6	410.5	33.6	343	10	US-09-948-094-2
7	408.5	33.5	453	9	US-09-978-295A-69
8	408.5	33.5	453	9	US-09-978-192A-69
9	408.5	33.5	453	9	US-09-978-192A-69
10	408.5	33.5	453	12	US-10-052-586-64
11	408	33.4	248	10	US-09-925-301-1017
12	408	33.4	327	10	US-09-804-156-16
13	408	33.4	327	10	US-09-946-633-8
14	408	33.4	454	10	US-09-888-615-103
15	397	32.5	791	9	US-09-967-386-1
16	397	32.5	810	10	US-09-946-893-2
17	389.5	31.9	812	9	US-09-335-325-1
18	389.5	31.9	812	10	US-09-788-142-1
19	389.5	31.9	812	10	US-09-761-120-1

ALIGNMENTS

RESULT 1
US-09-925-297-529 Application US/09925297
; Sequence 529, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 529
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-297-529

Query Match	81.6%	Score 996;	DB 10;	Length 192;
Best Local Similarity	99.5%;	Pred. No. 8.4e-83;		
Matches 189;	Conservative	0;	Mismatches 1;	Indels 0; Gaps 0;
Qy	5	SUTLSLVLLGSSWGGCIPAKPALSFSQRIYNGENAVLGSWFWQVSLQDSSGFHFCGSSL	54	
Db	1	SUTLSLVLLGSSWGGCIPAKPALSFSQRIYNGENAVLGSWFWQVSLQDSSGFHFCGSSL	60	
Qy	65	ISQSWVTAAHCNVSFGRHFVVLGVEYDRSSNAEPQLQVLSRAITHPSWNSTTMMNDVTL	124	
Db	61	ISQSWVTAAHCNVSFGRHFVVLGVEYDRSSNAEPQLQVLSRAITHPSWNSTTMMNDVTL	120	
Qy	125	LKLSPAQYTRISPVCLASSNEALTEGLTCVTTGWRSLSGVGNVTPAHLOQVALPLTV	184	
Db	121	LKLSPAQYTRISPVCLASSNEALTEGLTCVTTGWRSLSGVGNVTPAHLOQVALPLTV	180	
Qy	185	NOCROYWSSS	194	
Db	181	NOCROYWSSS	190	

Qy 231 I 231

Db 235 V 235

RESULT 5

US-09-925-297-695
Sequence 695, Application US/09925297
Patent No. US20020081659A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA105

CURRENT APPLICATION NUMBER: US/09/925,297

CURRENT FILING DATE: 2001-08-10

PRIOR APPLICATION NUMBER: PCT/US00/05989

PRIOR FILING DATE: 2000-03-08

PRIOR APPLICATION NUMBER: 60/124,270

PRIOR FILING DATE: 1999-03-12

NUMBER OF SEQ ID NOS: 928

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 695

LENGTH: 273

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (27)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (28)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (34)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-925-297-695

Query Match 34.4%; Score 420.5; DB 10; Length 273;
Best Local Similarity 40.7%; Pred. No. 9.2e-31;
Matches 98; Conservative 34; Mismatches 92; Indels 17; Gaps 8;

Qy 1 MLLSLTSLVLLGSSWCGGIPAKIPALSFQRIYNGENAVLGSWPPQVLSQ-DSSG--F 57

Db 5 MLRLSSLLLVAVASGYG-----PPSSXXSRVVGEDAVPYSPWPQVLSQYKSGSFY 58

Qy 58 HFCGSLISQSWVYTAACHNVSPGRHFVVLGEYDRSSNAEPQLVLSVS--RAITHPSWNS 115

Db 59 HFCGSLIAPQWVYTAGHCISRLDTYQVVLGEYNLAVKEGPQVPIINSEELFVHPLWNR 118

Qy 116 TTM--NNDVTLKLKSPAQYTRISPVCLASSNEALTEGLTCVTTGWGRLSGVGNVTPAH 173

Db 119 SCVACGNDIALIKLSRAQLGDAVOLASLPAGDILPNKTPCYITGWGRLYTNGLP-PDK 177

Qy 174 LQOVALPLVTNQCRQ--YWDSSITDSMICAGGASSCGSGGPLVC-QKGNFTWVLIG 230

Db 178 LQOARLPVVDYKHCSSRNWNGSVTKTWCAGGYIRSCNGDSGGPLNCPTEDGGQVHG 237

Qy 231 I 231

Db 238 V 238

RESULT 6

US-09-948-094-2

Sequence 2, Application US/09948094

Patent No. US20020090625A1

GENERAL INFORMATION:

APPLICANT: The Brigham and Women's Hospital, Inc.

APPLICANT: Mok, Samuel

APPLICANT: Wong, Kwong-kwok

TITLE OF INVENTION: Methods of Detecting Cancer Based on Prostatein

FILE REFERENCE: 81994/282423

CURRENT APPLICATION NUMBER: US/09/948,094

CURRENT FILING DATE: 2001-09-07

NUMBER OF SEQ ID NOS: 4

SOFTWARE: PatentIn version 3.0

SEQ ID NO 2

LENGTH: 343

TYPE: PRT

ORGANISM: Homo sapiens

US-09-948-094-2

Query Match 33.6%; Score 410.5; DB 10; Length 343;
Best Local Similarity 36.3%; Pred. No. 9.8e-30;
Matches 91; Conservative 38; Mismatches 89; Indels 33; Gaps 8;

Qy 4 LSLTSLVLLGSSWG-----CGIPAKIPALSFQRIYNGENAVLGSWPPQVLSQDSSG 56

Db 15 VAILLYLGLLRSGTGAEGAEAPCGVAP-----QARITGSSAVAGQMPQVSI-TYEG 66

Qy 57 HFCGSLISQSWVYTAACHNVSPGRHF-----VVLGEYDRSSNAEPQLVLSVRAITHP 111

Db 67 VHVCGSLVSEQWVLSAAHC--FPSEHHKEAYEVKLGHAQLDSYSEDKAVSTLKDIIPHP 124

Qy 112 SWNSTMMNDVTLKLKSPAQYTRISPVCLASSNEALTEGLTCVTTGWGRLS-GVGNVT 170

Db 125 SYLQEGSGDIALQLSRPITESRVIRPICLPAANASFPNGLHCTVTGHWVAPSLSLT 184

Qy 171 PAHQOVALPLVTNQCRQYWDSS-----ITDSMICAG--GAGASSCGDSGGPLVC 220

Db 185 PKPQLOEVLPLISRETCLYNIDAKPEHPFVQEDMVCAGYVEGKDKACDGSGLSC 244

Qy 221 QKGNFTWVLIGI 231

Db 245 PVEGLWYLTGI 255

RESULT 7

US-09-978-295A-69

Sequence 69, Application US/09978295A

Patent No. US20020156006A1

GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi

APPLICANT: Baker Kevin P.

APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan

APPLICANT: Ferrara, Napoleon

APPLICANT: Filvaroff, Ellen

APPLICANT: Fong, Sherman

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerber, Hanspeter

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, J. Christopher

APPLICANT: Gurney, Austin L.

APPLICANT: Hillan, Kenneth J.

APPLICANT: Kljavin, Ivar J.

APPLICANT: Kuo, Sophia S.

APPLICANT: Napier, Mary A.

APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann

APPLICANT: Shelton, David L.

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William I.

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

TITLE OF INVENTION: Acids Encoding the Same

FILE REFERENCE: P2630PIC11

CURRENT APPLICATION NUMBER: US/09/978,295A

CURRENT FILING DATE: 2001-10-15

PRIOR APPLICATION NUMBER: 09/918585

1 PRIOR FILING DATE: 2001-07-30
2 PRIOR APPLICATION NUMBER: 60/062250
3 PRIOR FILING DATE: 1997-10-17
4 PRIOR APPLICATION NUMBER: 60/064249
5 PRIOR FILING DATE: 1997-11-03
6 PRIOR APPLICATION NUMBER: 60/065311
7 PRIOR FILING DATE: 1997-11-13
8 PRIOR APPLICATION NUMBER: 60/066364
9 PRIOR FILING DATE: 1997-11-21
10 PRIOR APPLICATION NUMBER: 60/077450
11 PRIOR FILING DATE: 1998-03-10
12 PRIOR APPLICATION NUMBER: 60/077632
13 PRIOR FILING DATE: 1998-03-11
14 PRIOR APPLICATION NUMBER: 60/077641
15 PRIOR FILING DATE: 1998-03-11
16 PRIOR APPLICATION NUMBER: 60/077649
17 PRIOR FILING DATE: 1998-03-11
18 PRIOR APPLICATION NUMBER: 60/077791
19 PRIOR FILING DATE: 1998-03-12
20 PRIOR APPLICATION NUMBER: 60/078004
21 PRIOR FILING DATE: 1998-03-13
22 PRIOR APPLICATION NUMBER: 60/078886
23 PRIOR FILING DATE: 1998-03-20
24 PRIOR APPLICATION NUMBER: 60/078936
25 PRIOR FILING DATE: 1998-03-20
26 PRIOR APPLICATION NUMBER: 60/078910
27 PRIOR FILING DATE: 1998-03-20
28 PRIOR APPLICATION NUMBER: 60/078939
29 PRIOR FILING DATE: 1998-03-20
30 PRIOR APPLICATION NUMBER: 60/079294
31 PRIOR FILING DATE: 1998-03-25
32 PRIOR APPLICATION NUMBER: 60/079656
33 PRIOR FILING DATE: 1998-03-26
34 PRIOR APPLICATION NUMBER: 60/079664
35 PRIOR FILING DATE: 1998-03-27
36 PRIOR APPLICATION NUMBER: 60/079689
37 PRIOR FILING DATE: 1998-03-27
38 PRIOR APPLICATION NUMBER: 60/079663
39 PRIOR FILING DATE: 1998-03-27
40 PRIOR APPLICATION NUMBER: 60/079728
41 PRIOR FILING DATE: 1998-03-27
42 PRIOR APPLICATION NUMBER: 60/079786
43 PRIOR FILING DATE: 1998-03-27
44 PRIOR APPLICATION NUMBER: 60/079920
45 PRIOR FILING DATE: 1998-03-30
46 PRIOR APPLICATION NUMBER: 60/079923
47 PRIOR FILING DATE: 1998-03-30
48 PRIOR APPLICATION NUMBER: 60/080105
49 PRIOR FILING DATE: 1998-03-31
50 PRIOR APPLICATION NUMBER: 60/080107
51 PRIOR FILING DATE: 1998-03-31
52 PRIOR APPLICATION NUMBER: 60/080165
53 PRIOR FILING DATE: 1998-03-31
54 PRIOR APPLICATION NUMBER: 60/080194
55 PRIOR FILING DATE: 1998-03-31
56 PRIOR APPLICATION NUMBER: 60/080327
57 PRIOR FILING DATE: 1998-04-01
58 PRIOR APPLICATION NUMBER: 60/080328
59 PRIOR FILING DATE: 1998-04-01
60 PRIOR APPLICATION NUMBER: 60/080333
61 PRIOR FILING DATE: 1998-04-01
62 PRIOR APPLICATION NUMBER: 60/080334
63 PRIOR FILING DATE: 1998-04-01
64 PRIOR APPLICATION NUMBER: 60/081070
65 PRIOR FILING DATE: 1998-04-08
66 PRIOR APPLICATION NUMBER: 60/081049
67 PRIOR FILING DATE: 1998-04-08
68 PRIOR APPLICATION NUMBER: 60/081071
69 PRIOR FILING DATE: 1998-04-08
70 PRIOR APPLICATION NUMBER: 60/081195
71 PRIOR FILING DATE: 1998-04-08
72 PRIOR APPLICATION NUMBER: 60/081203
73 PRIOR FILING DATE: 1998-04-09
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000


```

; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match          33.5%: Score 408.5; DB 9: Length 453;
Best Local Similarity 39.9%: Pred No. 2,le-29;
Matches 83; Conservative 31; Mismatches 83; Indels 11; Gaps 6;

QY 30 FQRTVNGENAVLGSWPMQVSLQDSGFHFCGSLISQSWVYTAACHNVS---PGRHFVY 86
   : |||| : : |||| |||| : |||| |||| : |||| |||| : |||| |||| :
Db 213 YSSRIVGGNMSLSQWPQASLQ-EGYHLCGGSVITPLWIITAAHCYVDLYLPKSWTIQ 271

QY 87 LGEYDRSSNAEPLOVLSVSRATHPSWNSWTNNVDVTLKLTASPAQYTRISPVCLASN 146
   : || : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 272 VGLVSLLDNPAFSHL--VEKIYVSKYKPKRLGNDIALMKLPLTFNEMIQVCLPNS 329

QY 147 EALTEGLTCVTGMRGLSGVGNVTAHLQOVALPLVTYNOC--ROYWSSITDMSICAG- 203
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 330 ENFPDGKVCWTSWCATGDDGASPV-LNHAAPLISKNKICHRVDYGGIISPSMLCAGY 388

QY 204 -GAGASSCGDGGGGLVPCQKGNFTWVLIG 230
   | ||||| ||||| ||||| : : : : : : : : : : : : : : : : : :
Db 389 LTGVDVSCQDGGGGLVPCQERRLKLVG 416

RESULT 8
US-09-978-697-69
; Sequence 69, Application US/09978697
; Patent No. US20020169284A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrata, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PlC27
; CURRENT APPLICATION NUMBER: US/09/978,697
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
```


APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C9
CURRENT APPLICATION NUMBER: US/09/978,192A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 05/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598

; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 33.5%; Score 408.5; DB 9; Length 453;
Best Local Similarity 39.9%; Pred. No. 2.1e-29;
Matches 83; Conservative 31; Mismatches 83; Indels 11; Gaps 6;

QY 30 FSRIVNGENAVLGSWPQVSLQDSGFHFCGSLISQSWVTAHCNVS---PGRHFV 86
DB 213 YSRIVGGMNLSLSPQWQASLQ-FQYHLCGGSVITPLWITAAHCNVYDLYLPKSWTIQ 271

QY 87 LGEYDRSSNAELQVLSVSRATHPSWNTTMNDVTLKLASPAQYTRISPCVCLASSN 146
DB 272 VGLVSLDNPAPSHL--VEKIVYHSKYKPKRLGNDIALMKLAGPLTFNEMIQVCLPNSE 329

QY 147 EALTEGLCTVTGWRGLSGVGNVTPAHLQOVALPLVTVNO--RQWDSSTIDSMTCAG- 203
DB 330 ENFPDGKVCWTSGWATEDGGDASPY-LNHAAPVLISNKNICHRDVGYGIIISPMLCAGY 388

QY 204 -GAGASSCOGDSGGLVLCOKGNTWVLIG 230
DB 389 LTGGVDSOCGDSGGLVLCQERRLKLVG 416

RESULT 10
US-10-052-586-64
; Sequence 64, Application US/10052586
; Patent No. US20020127584A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C1
; CURRENT APPLICATION NUMBER: US/10/052,586
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18

; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063564
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063734
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063870
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066120
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066466
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069335
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069425
; PRIOR FILING DATE: 1997-12-12
; PRIOR APPLICATION NUMBER: 60/069870
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/068017
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569

```

RESULT 11
US-09-925-301-1017
; Sequence 1017, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925, 301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCI/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124, 270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1017
; LENGTH: 248

```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1017

Query Match      33.4%; Score 408; DB 10; Length 248;
Best Local Similarity 40.2%; Pred. No. 1.1e-29;
Matches 84; Conservative 31; Mismatches 82; Indels 12; Gaps 7;

QY 30 FSRIVGENAVLGSWPMQVSLQDSSGFHFCGSLISQSWVYTAACNVSV---PGRHFVW 86
Db 7 YSRIVGGMNLSLQWPMQASLQ-FQGYHLGCGSVITPLWIIITAAHCYVDLYLPKSWTIQ 65

QY 87 LGEYDRSSNAEPLQVLSVRAITHPSWNSMTNMNDVTLKLKASPAQYTRISPVCLASN 146
Db 66 VGLVSLLDNPAPSHL--VEKIVYHSKYKPKRLGNDIALMKLAGPLTFNEMIQPVCPLPSE 123

QY 147 EALTEGLTCVTGTGRL--SGVGNVTPAHQQVALPLVTNOC--ROYWDSSTIDSMICAG 203
Db 124 ENFPDGKVCWTSGWATEDGAGDASPV-LNHAAPVPLISNKNICNHRDVGIIISPSMLCAG 182

QY 204 --GAGASSCGDGGPLVCOKGNTWVLIG 230
Db 183 YLTGGVDSQDGGGGLVCQERRLKLVG 211

RESULT 12
US-09-804-156-16
; Sequence 16, Application US/09804156
; Patent No. US20020068320A1
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: Serine protease Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PT00594
; CURRENT APPLICATION NUMBER: US/09/804,156
; CURRENT FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/189,025
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-804-156-16

Query Match      33.4%; Score 408; DB 10; Length 327;
Best Local Similarity 40.2%; Pred. No. 1.5e-29;
Matches 84; Conservative 31; Mismatches 82; Indels 12; Gaps 7;

QY 30 FSRIVGENAVLGSWPMQVSLQDSSGFHFCGSLISQSWVYTAACNVSV---PGRHFVW 86
Db 86 YSRIVGGMNLSLQWPMQASLQ-FQGYHLGCGSVITPLWIIITAAHCYVDLYLPKSWTIQ 144

QY 87 LGEYDRSSNAEPLQVLSVRAITHPSWNSMTNMNDVTLKLKASPAQYTRISPVCLASN 146
Db 145 VGLVSLLDNPAPSHL--VEKIVYHSKYKPKRLGNDIALMKLAGPLTFNEMIQPVCPLPSE 202

QY 147 EALTEGLTCVTGTGRL--SGVGNVTPAHQQVALPLVTNOC--ROYWDSSTIDSMICAG 203
Db 203 ENFPDGKVCWTSGWATEDGAGDASPV-LNHAAPVPLISNKNICNHRDVGIIISPSMLCAG 261

QY 147 EALTEGLTCVTGTGRL--SGVGNVTPAHQQVALPLVTNOC--ROYWDSSTIDSMICAG 203
Db 203 ENFPDGKVCWTSGWATEDGAGDASPV-LNHAAPVPLISNKNICNHRDVGIIISPSMLCAG 261

QY 204 --GAGASSCGDGGPLVCOKGNTWVLIG 230
Db 262 YLTGGVDSQDGGGGLVCQERRLKLVG 290

RESULT 13
US-09-946-633-8
; Sequence 8, Application US/09946633
; Patent No. US20020119925A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: PT005P1
```

```
; FILE REFERENCE: Serine proteases
; CURRENT APPLICATION NUMBER: US/09/946,633
; CURRENT FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 09/597,839
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 60/133,239
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 60/135,163
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 60/147,005
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/152,935
; PRIOR FILING DATE: 1999-09-09
; PRIOR APPLICATION NUMBER: 60/162,979
; PRIOR FILING DATE: 1999-11-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-946-633-8

Query Match      33.4%; Score 408; DB 10; Length 327;
Best Local Similarity 40.2%; Pred. No. 1.5e-29;
Matches 84; Conservative 31; Mismatches 82; Indels 12; Gaps 7;

QY 30 FSRIVGENAVLGSWPMQVSLQDSSGFHFCGSLISQSWVYTAACNVSV---PGRHFVW 86
Db 86 YSRIVGGMNLSLQWPMQASLQ-FQGYHLGCGSVITPLWIIITAAHCYVDLYLPKSWTIQ 144

QY 87 LGEYDRSSNAEPLQVLSVRAITHPSWNSMTNMNDVTLKLKASPAQYTRISPVCLASN 146
Db 145 VGLVSLLDNPAPSHL--VEKIVYHSKYKPKRLGNDIALMKLAGPLTFNEMIQPVCPLPSE 202

QY 147 EALTEGLTCVTGTGRL--SGVGNVTPAHQQVALPLVTNOC--ROYWDSSTIDSMICAG 203
Db 203 ENFPDGKVCWTSGWATEDGAGDASPV-LNHAAPVPLISNKNICNHRDVGIIISPSMLCAG 261

QY 204 --GAGASSCGDGGPLVCOKGNTWVLIG 230
Db 262 YLTGGVDSQDGGGGLVCQERRLKLVG 290

RESULT 14
US-09-888-615-103
; Sequence 103, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 103
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-103

Query Match      33.4%; Score 408; DB 10; Length 454;
Best Local Similarity 40.2%; Pred. No. 2.3e-29;
Matches 84; Conservative 31; Mismatches 82; Indels 12; Gaps 7;
```

```
Qy 30 FSQRIVNGENAVLGSWPQVSLQSDSGFHFCGGLSQSQQSWYVTAHNCVSV---PGRHFV 86
Db 213 YSSRIVGNGNLLSQWPMQASLQ-FQGYHLCCGGVITPLWITTAARCVYDLYLPKSWTIQ 271
Qy 87 LGEYDRSSNAEPLQVLSVSRATHPSWNSNTMNDVTLKLLASPAQYTTTRISPVCLASSN 146
Db 272 VGLVSLLDNPAPSHL--VEKIVYHSKYKPKRLGNDIALMKLAGPLTFNEMIQVCLPNS 329
Qy 147 EALTEGLTCVTTGRL--SGVGNVTPAHLQVALPLVTNOC--ROYWSSITDSMICAG 203
Db 330 ENFPDGKVCWTSNGATEDGAGDASPV-LNHAAPLISNKICNHRDVIYGGIISFSLCAG 388
Qy 204 --GAGASSCQCGSGPLVCQKGNWTWVLIG 230
Db 389 YLTGGVDSQCGSGPLVCQERRLWKLVG 417
```

RESULT 15

```
US-09-967-386-1
; Sequence 1, Application US/09967386
; Patent No. US2002015992A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Henkin, Jack
; APPLICANT: Davidson, Donald J.
; TITLE OF INVENTION: ANTIANGIOGENIC POLYPEPTIDES AND METHODS
; FILE REFERENCE: 6738.US.02
; CURRENT APPLICATION NUMBER: US/09/967,386
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/236,550
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 791
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-967-386-1
```

```
Query Match      32.5%; Score 397; DB 9; Length 791;
Best Local Similarity 39.3%; Pred. No. 4.6e-28;
Matches 88; Conservative 35; Mismatches 83; Indels 18; Gaps 8;

Qy 16 SWGGGIPAIKPAKLSFQRIVNGENAVLGSWPQVSLQSDSGFHFCGGLSQSQQSWYVTAAH 75
Db 545 SFDCGKQVEPK-KCPGRVGGCVAHPSHPWQVSLRFRGMHFCGGTLLSPWVLTAAH 603

Qy 76 C---NVSPGRHFVVLGEYDRSSNAEP-LQVLSVSRATHPSWNSNTMNDVTLKLLASPA 131
Db 604 CLEKSPRPSSYKVLGAH-QEVNLEPHVQETEVSRLEPT-----RKDIALKLLSSPA 656

Qy 132 QYTRISPVCLASSNEALTEGLTCVTGWRGLSGVGNVTPAHLQVALPLVTNOCROY- 190
Db 657 VITDKVIPACLPSPNYVADRTECFITGWGETQGTFGA--GLLKEAQLPVIENKVCNRYE 714

Qy 191 -WSSSITDSMICAG--GAGASSCQCGSGPLVCQKGNWTWVLIG 231
Db 715 FLNGRVSTELCAGHLAGGTDCQCGSGPLVCQERRLWKLVG 758
```

Search completed: December 20, 2002, 15:21:06
Job time : 8.30435 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:14:08 ; Search time 28.6087 Seconds
(without alignments)
1663.721 Million cell updates/sec

Title: US-09-856-319B-2_COPY_1_231

Perfect score: 1221

Sequence: 1 MLLLSLTSLVLLSSWGCG.....GDSGGPLVCQKGNWTWVLGI 231

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1102	90.3	264	11 Q9EQZ8	Q9eqz8 rattus norv
2	1093	89.5	264	11 Q9D7P8	Q9d7p8 mus musculu
3	1093	89.5	264	11 Q9ER05	Q9er05 mus musculu
4	1083	88.7	264	11 Q9D960	Q9d960 mus musculu
5	766.5	62.8	261	13 Q9W7Q4	Q9w7q4 paralichthy
6	710.5	58.2	260	13 Q9W7Q3	Q9w7q3 paralichthy
7	684.5	56.1	263	11 Q9DC86	Q9dc86 mus musculu
8	682.5	55.9	263	11 Q9CR35	Q9cr35 mus musculu
9	676.5	55.4	263	11 Q9D8X8	Q9d8x8 mus musculu
10	674.5	55.2	263	13 Q9PWQ6	Q9pwq6 gadus morhu
11	472.5	38.7	269	11 Q9C052	Q9cq52 mus musculu
12	465.5	38.1	266	13 Q92077	Q92077 gadus morhu
13	464.5	38.0	269	11 Q9D7T9	Q9d7t9 mus musculu
14	454.5	37.2	260	13 Q9W7P9	Q9w7p9 paralichthy
15	447	36.6	164	11 Q9DC82	Q9dc82 mus musculu
16	442	36.2	249	13 Q9W7Q1	Q9w7q1 paralichthy

17	434	35.5	269	4 Q96QV5	Q96qv5 homo sapien
18	423	34.6	267	5 Q9BK47	Q9bk47 luidia foli
19	422	34.6	270	13 Q91039	Q91039 gadus morhu
20	420.5	34.4	270	4 Q96QL8	Q96ql8 homo sapien
21	420	34.4	277	5 Q96899	Q96899 scolopendra
22	417.5	34.2	266	13 Q9W7Q0	Q9w7q0 paralichthy
23	417.5	34.2	269	6 Q95KW7	Q95kw7 bos taurus
24	411	33.7	268	13 Q9W7Q2	Q9w7q2 paralichthy
25	409	33.5	1524	13 Q91674	Q91674 xenopus lae
26	400.5	32.8	257	6 Q19023	Q19023 macaca mula
27	400.5	32.8	339	11 Q99144	Q99144 mus musculu
28	399.5	32.7	453	11 Q8VDE0	Q8vd0 mus musculu
29	399	32.7	263	5 Q62562	Q62562 penaeus van
30	398.5	32.6	264	13 Q8QGF6	Q8qgf6 xenopus lae
31	398	32.6	810	4 Q15146	Q15146 homo sapien
32	392.5	32.1	812	11 Q91WJ5	Q91wj5 mus musculu
33	392	32.1	334	6 Q46507	Q46507 papio hamad
34	389.5	31.9	247	13 Q42608	Q42608 petromyzon
35	389.5	31.9	767	13 Q9DGR2	Q9dgr2 xenopus lae
36	388.5	31.8	247	13 Q42158	Q42158 petromyzon
37	388.5	31.8	331	11 Q8R1A6	Q8rla6 mus musculu
38	386	31.6	787	5 Q9VEY6	Q9vey6 drosophila
39	386	31.6	1004	13 P79953	P79953 xenopus lae
40	385.5	31.6	273	11 Q921N4	Q921n4 mus musculu
41	385.5	31.6	812	11 Q9ROW3	Q9row3 rattus norv
42	381	31.2	266	6 Q46644	Q46644 macaca fasc
43	381	31.2	266	11 Q91X79	Q91x79 mus musculu
44	380	31.1	454	6 Q46506	Q46506 papio hamad
45	379	31.0	492	4 Q96T73	Q96t73 homo sapien

ALIGNMENTS

RESULT 1

Q9EQZ8 ID Q9EQZ8 PRELIMINARY; PRT; 264 AA.
AC Q9EQZ8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Chymopasin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=RAT PANCREAS;
RA Sogame Y., Mitsui S., Kataoka K., Kashima K., Kato M., Sakagami J.,
RA Yanaguchi N.;
RT "Molecular cloning of rat chymopasin.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB020757; BAB20287.1; -;
DR HSSP; P00766; 4CHA.
DR MEROPS; S01.256; -;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Sec; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28116 MW; F9ED5D210FD3500E CRC64;

Query Match 90.3%; Score 1102; DB 11; Length 264;
Best Local Similarity 87.0%; Pred. No. 5.1e-92;
Matches 201; Conservative 17; Mismatches 13; Indels 0; Gaps 0;

```

QY 1 MLLSLTSLVLGSSWGCGIPAIPALFSORIVNGENAVLGSWPQVSLQDSSGFHFC 60
Db 1 MLLSLTSLVLGSSWGCGIPAIPALFSORIVNGENAVLGSWPQVSLQDNTGFHFC 60
QY 61 GGLISQSWVVTAAHCNVSPGRHFVVLGEYDRSSNAEPQLVLSVRAITHPSWNSTMMN 120
Db 61 GGLIAPNWWVTAAHCNVTPGRHFVVLGEYDRSSNAEPQLVLSVRAITHPSWNSTMMN 120
QY 121 DVTLKLKASPAQYTTTRISPVCLASSNEALTEGLTCVTTGWRSLSGVGNVTPAHLQOVALP 180
Db 121 DLTLKLKASPARYTAQVSPVCLASNEALPAGLTCTVTTGWRSLSGVGNVTPAHLQOVALP 180
QY 181 LVTVNOCROYWDSITDSMICAGGASSCGQSGGGLVCQKGNWTWVLIGI 231
Db 181 LVTVNOCROYWGSRTDSMICAGGASSCGQSGGGLVCQKGNWTWVLIGI 231

RESULT 2
Q9D7P8 PRELIMINARY; PRT; 264 AA.
AC Q9D7P8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE 1810004D15rik protein.
GN CTRL OR 1810004D15RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT Functional annotation of a full-length mouse cDNA collection.;
RL Nature 409:685-690(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AK009019; BAB26029.1; -.
DR HSSP; P00766; 4CHA.
DR MEROPS; S01.256; -.
DR MGD; MGI:88558; Ctrl.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28167 MW; 1D979469A07056C2 CRC64;

Query Match 89.5%; Score 1093; DB 11; Length 264;
Best Local Similarity 85.7%; Pred. No. 3.3e-91;
Matches 198; Conservative 21; Mismatches 12; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGSSWGCGIPAIPALFSORIVNGENAVLGSWPQVSLQDSSGFHFC 60
Db 1 MLLSLTSLVLGSSWGCGIPAIPALFSORIVNGENAVLGSWPQVSLQDNTGFHFC 60
QY 61 GGLISQSWVVTAAHCNVSPGRHFVVLGEYDRSSNAEPQLVLSVRAITHPSWNSTMMN 120
Db 61 GGLIAPNWWVTAAHCNVTPGRHFVVLGEYDRSSNAEPQLVLSVRAITHPSWNSTMMN 120
QY 121 DVTLKLKASPAQYTTTRISPVCLASSNEALTEGLTCVTTGWRSLSGVGNVTPAHLQOVALP 180
Db 121 DLTLKLKASPARYTAQVSPVCLASNEALPAGLTCTVTTGWRSLSGVGNVTPAHLQOVALP 180
QY 181 LVTVNOCROYWDSITDSMICAGGASSCGQSGGGLVCQKGNWTWVLIGI 231
Db 181 LVTVNOCROYWGSRTDSMICAGGASSCGQSGGGLVCQKGNWTWVLIGI 231

RESULT 3
Q9ER05 PRELIMINARY; PRT; 264 AA.
AC Q9ER05;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Chymopasin (Chymotrypsin A CTra-1).
GN CTRL OR CTra1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Mitsui S., Yamaguchi N.;
RT "Molecular cloning of mouse chymopasin.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=12956/SVEVTA; TISSUE=SPLEEN;
RA Bjoernlett M.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB016228; BAB20275.1; -.
DR EMBL; AF236365; AAL11034.1; -.
DR HSSP; P00766; 4CHA.
DR MEROPS; S01.256; -.
DR MGD; MGI:88558; Ctrl.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28135 MW; 1D979709A07056C2 CRC64;

Query Match 89.5%; Score 1093; DB 11; Length 264;
Best Local Similarity 85.7%; Pred. No. 3.3e-91;
Matches 198; Conservative 21; Mismatches 12; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGSSWGCGIPAIPALFSORIVNGENAVLGSWPQVSLQDSSGFHFC 60
Db 1 MLLSLTSLVLGSSWGCGIPAIPALFSORIVNGENAVLGSWPQVSLQDNTGFHFC 60
QY 61 GGLISQSWVVTAAHCNVSPGRHFVVLGEYDRSSNAEPQLVLSVRAITHPSWNSTMMN 120
Db 61 GGLIAPNWWVTAAHCNVTPGRHFVVLGEYDRSSNAEPQLVLSVRAITHPSWNSTMMN 120
QY 121 DVTLKLKASPAQYTTTRISPVCLASSNEALTEGLTCVTTGWRSLSGVGNVTPAHLQOVALP 180
Db 121 DLTLKLKASPARYTAQVSPVCLASNEALPAGLTCTVTTGWRSLSGVGNVTPAHLQOVALP 180
QY 181 LVTVNOCROYWDSITDSMICAGGASSCGQSGGGLVCQKGNWTWVLIGI 231
Db 181 LVTVNOCROYWGSRTDSMICAGGASSCGQSGGGLVCQKGNWTWVLIGI 231

Query Match 89.5%; Score 1093; DB 11; Length 264;
Best Local Similarity 85.7%; Pred. No. 3.3e-91;
Matches 198; Conservative 21; Mismatches 12; Indels 0; Gaps 0;

```

Db	121	DLTLLKASPARYAQSPVCLASTNEALPSGUTCTVTTGWRISGVGNTPARLQOVL	180
Qy	181	LVTVNCRQYWDSSITDSMICAGAGASSCGQSGGGLVCQKGNVTWLIGI	231
Db	181	LVTVNCRQYWGARIITDAMICAGSGASSCGQSGGGLVCQKGNVTWLIGI	231
RESULT 4			
Q90960			
ID	Q9D960	PRELIMINARY; PRT; 264 AA.	
AC	Q9D960;		
DT	01-JUN-2001 (T-EMBLrel. 17, Created)		
DT	01-JUN-2001 (T-EMBLrel. 17, Last sequence update)		
DT	01-JUN-2002 (T-EMBLrel. 21, Last annotation update)		
DE	1810004D15Rik protein.		
GN	1810004D15Rik		
OS	Mus musculus (Mouse)		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=PANCREAS;		
RC	MEDLINE=21085660; PubMed=11217851;		
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,		
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,		
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Asnburner M., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Glassi C., King B., Kochiwa H.,		
RA	Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,		
RA	Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,		
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,		
RA	Brownstein M.J., Bult C., Fietcher C., Fujita M., Gariboldi M.,		
RA	Gustinchin S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,		
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,		
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,		
RA	Suzuki H., Toyooka K., Wang K.H., Weltz C., Whittaker C., Wilming L.,		
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kotsuki S.,		
RA	Hayashizaki Y.		
RT	Functional annotation of a full-length mouse cdna collection.;		
RL	Nature 409:685-690(2001).		
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.		
DR	EMBL: AK007333; BAB24967.1; -.		
DR	HSSP: P00766; 4CHA.		
DR	MEROPS: S01.256; -.		
DR	MGI: 88558; Ctrl.		
DR	InterPro: IPR001314; Chymotrypsin.		
DR	InterPro: IPR001254; Ser-protease_Try.		
DR	Pfam: P00089; trypsin.1.		
DR	PRINTS: PR00722; CHYMOTRYPSIN.		
DR	SMART: SM00020; Tryp_Spc.1.		
DR	PROSITE: PS50240; TRYPSIN_DOM; 1.		
DR	PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.		
DR	PROSITE: PS00135; TRYPSIN_SER; 1.		
DR	Hydrolase; Serine protease.		
SW	SEQUENCE 264 AA; 28151 MW; 1D979719E07C16DE CRC64;		

	Query Match	88.7%	Score 1083;	DB 11;	Length 264;
	Best Local Similarity	85.3%;	Pred. No. 2.7e-90;		
	Matches 197;	Conservative 21;	Mismatches 13;	Indels 0;	Gaps 0;
Qy	1	MLLSLTLSLVLLGSSNGCGIPATKPLSPFSQRTIVNGENAVLGSWPQVSLQDSSGGPHFC	60		
Db	1	MLLSLTLTLVLLGSSNGCGIPATPALSYNQRTIVNGENAVPGSWPQVSLQDNTGPHFC	60		
Qy	61	GGSLISGSWVYTAACHNVSFCGRHFVVLGEYDRSSNAPQLQVLVSRAITHPSWNSWTMNN	120		
Db	61	GGSLISPNWVYTAACHQVTPGRHFVVLGEYDRSSNAPQVLSTLARAITHPNWNAITMNN	120		

Qy	121	DVTLKLLKASPAQYTRTRISPCVLASSNEALTEGILTCVTTGMRGSLSGVGNVTPAHLLQOQVALP	180
Db	121	DLTLKLLKASPARYTAQYTRTRISPCVLASSNEALPSGILTCVTTGMRISGSGVGNVTPARLQOQVLP	180
Qy	181	LVTYNQCRQWDSISITDSMTICAGAGASSCGDSGGGLPVCQKGNWTWYLYGI	231
Db	181	LVTYNQCRQWYGARITDAMTCAGSGASSCGDSGGGLPVCQKGNWTWYLYGI	231

```

RESULT 5
Q9W7Q4
ID Q9W7Q4 PRELIMINARY; PRT; 261 AA.
AC Q9W7Q4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chymotrypsinogen 1.
OS Paralicthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidae; Paralicthyidae; Paralicthys.
NCBI_TaxID=8255;
RX [1]
RN SEQUENCE FROM N. A.
RP TISSUE=PANCREAS;
RC Suzuki T., Srivastava A.S., Kurokawa T.;
RA "Japanese flounder mRNA for chymotrypsinogen 1.";
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB029753; BAA82365.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.256; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR HYDROLASE; Serine protease.
SQ SEQUENCE 261 AA; 28184 MW; D7090A9D65395B7D CRC64;

```

Query Match	62.8%	Score	766.5	DB 13	Length	261	
Best Local Similarity	61.6%	Pred. No.	1.3e-61				
Matches	143	Conservative	31	Mismatches	53	Gaps	3

Qy	1	MLLLSLTSLVLVLLGSSWCCGIPAIKPALSPFSQRTVNGENAVLGSWPQVSLQDSSGPHFC	60
		: : : : : :	
Db	1	MLLFISCFALV--ASALGCGVPSIKPQVSGYKNVGETAVSGSWPQVSLQDGRGPHFC	58
Qy	61	GGSLISQWVYTAACHNVSQGRHFVLVGEYDRSSNAEPLQVLSVRAITHPSWSTTMNN	120
Db	59	GGSLISPVWVYTAACHTVSPNRHVRILGEHSDRYNNEFIQVMASTARATHPYYSNQFN	118
Qy	121	DVTLKLASPAQYTRISPVCLASNEALTGLTCVTTGWGRSLGCVGNVTPAHLQOQVALP	180
Db	119	DITLLRLSSPQMTSRVSPVCLASSTISPGTKCVTTGWGRTGOTS--SPRYLQOTSLP	176
Qy	181	LVTVNOCROYND--SITDSMTCAGAGASSCGDGGPLVCQKGNWTWVLCI	231
		: : : :	
Db	177	LLSPAQCQYGYGNRITDAMICAGASVSCQDGGPLVCEKSGAWFLTGI	228

RESULT	6
Q9W7Q3	
ID	Q9W7Q3
AC	Q9W7Q3
DT	01-NOV-1999 (TrEMBLrel. 12, Created)
DT	01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

```
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Chymotrypsinogen 2.
OS Paralicthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidi; Paralichthyidae; Paralichthys.
OX NCBI_TaxID=8255;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Suzuki T., Srivastava A.S., Kurokawa T.;
RT "Japanese flounder mRNA for chymotrypsinogen 2.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: AB029754; BAA82366.1; -.
DR HSSP: P00766; 1CHG.
DR MEROPS: S01.152; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR Hydrolase; Serine protease.
KW SEQUENCE 260 AA; 27793 MW; 9F583044E22F78C0 CRC64;

Query Match 58.2%; Score 710.5; DB 13; Length 260;
Best Local Similarity 57.7%; Pred. No. 1.6e-56;
Matches 128; Conservative 36; Mismatches 57; Indels 1; Gaps 1;

QY 10 LVLLGSGWGGIPAIKPAISFQRIYNGENAVLGSWPMQVSLQDSSGFHFCGSGSLISQSW 69
DB 7 LAFAGRAAGCGSPAIIPVYGYRIYNGEALPHSPWQVSLQDYTFHFCGSLINENW 66

QY 70 VVTAHCNVSFGRHVVYLGEDRSSNAEPLQVLSVSRATHPSWNSMTTNNVTLKLKAS 129
DB 67 VVTAHCNVRTS-HRVILGHDHSSAEDIQVMKVKFKPRYNGTYTNNIDILLIKLAA 125

QY 130 PAQYTRISPVCLASNEALTEGLTCVTTGWRSLGSGVGNVTPAHLQOVALPLVTVNQCRQ 189
DB 126 PAQNMNRVSPVCVAETSDNFAGMKCVTSGLTRHNPDPPTPALQOAAALPLTNDCCR 185

QY 190 YWDSSTITDSMICAGGAGASSCGDGGPLVCQKGNWTLVGI 231
DB 186 YWGNKISLNMICAGAGASSCGDGGPLVCQKAGAWTLVGI 227

RESULT 7
Q9DC86 ID Q9DC86 PRELIMINARY; PRT; 263 AA.
AC Q9DC86
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE 220000809Rik protein.
GN 220000809Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=SPLEEN;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
```

```
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Futuro M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann J., Mashima J., Mazzarelli J., Mombaerts P.,
RA Lyons P., Marchionni L., Mashima J., Mazzaresi I., Sakamoto N.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RA Functional annotation of a full-length mouse cDNA collection.";
RT Nature 403:685-690(2001).
RL -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: AK003060; BAB22539.1; -.
DR HSSP: P00766; 1GCT.
DR MEROPS: S01.152; -.
DR MGD; MGI:1913723; 220000809Rik.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 263 AA; 27821 MW; 2620A27AFBA5D04D CRC64;

Query Match 56.1%; Score 684.5; DB 11; Length 263;
Best Local Similarity 54.5%; Pred. No. 3.7e-54;
Matches 126; Conservative 38; Mismatches 66; Indels 1; Gaps 1;

QY 1 MLLLSITLSLVLLGSGWGGIPAIKPAISFQRIYNGENAVLGSWPMQVSLQDSSGFHFC 60
DB 1 MAFLMLVSCFALVATFGCGVPAIQVLTGLSRVNGEDAIPGSPWQVSLQDRTGFHFC 60

QY 61 GGLISQSVMVTAHCNVSFGRHVVYLGEDRSSNAEPLQVLSVSRATHPSWNSMTMNN 120
DB 61 GGLISENVVTAHCNVSFGRHVVYLGEDRSSNAEPLQVLSVSRATHPSWNSMTMNN 119

QY 121 DVTLLKASPAQYTRISPVCLASNEALTEGLTCVTTGWRSLGSGVGNVTPAHLQOVALP 180
DB 120 DITLLKLATPAQFSETVSACVLTPTVDDFFPAGTLCATTGWGKTKYNALKTDPKLOQAALP 179

QY 181 LVTVNCROYWDSITDSMICAGGAGASSCGDGGPLVCQKGNWTLVGI 231
DB 180 IVSEAKCKESWGSKITDVMICAGAGVSSCGMDGGPLVCQKQGVWTLVGI 230

RESULT 8
Q9CR35 ID Q9CR35 PRELIMINARY; PRT; 263 AA.
AC Q9CR35;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE 220000809Rik protein.
GN 220000809Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=STOMACH, SPLEEN, AND PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Aizawa K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
```

```
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., Fujita M., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC EMBL: AK008927; BAB25971.1; -.
DR EMBL: AK003079; BAB22553.1; -.
DR EMBL: AK007765; BAB25241.1; -.
DR EMBL: AK007815; BAB25280.1; -.
DR EMBL: AK008729; BAB25861.1; -.
DR EMBL: AK008888; BAB25954.1; -.
DR HSSP: P00766; IGCT.
DR MEROPS: S01.152; -.
DR MGD: MGI:1913723; 2200008D09RIK.
DR InterPro: IPR001314; Chymotrypsin.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PS02040; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 263 AA; 27822 MW; 28C4487AF1A26B27 CRC64;

Query Match 55.9%; Score 682.5; DB 11; Length 263;
Best Local Similarity 54.5%; Pred. No. 5.6e-54;
Matches 126; Conservative 37; Mismatches 67; Indels 1; Gaps 1;

QY 1 MLLSLTSLVLGGSGGCGIPALPSQRIYNGENAVLGSWPQVSLQDSSGFHFC 60
Db 1 MAFLWLVSFCALVGATFGCGVPAIQVLTGLSRIVNGEDAIPGSWPQVSLQDRTGFHFC 60
QY 61 GGSLSQSQSVVTAACHNVSFGRHFVVLGVEYDRSSNAEPLQVLSVSRATHPSWNSTTMN 120
Db 61 GGSLSSENWVTAACHGCVKT-TDVVVAGFDDQGSDEENVQVLKIAQVFNKFNSTFVRN 119
QY 121 DVTLLKASPAQYTRTRISPVCLASNEALTEGLTCVTTGWRGLSGVGNVTPAHLOQVALP 180
Db 120 DITLLKLATPAQFSETVSACVLTVDVDDPPATGLCATGTGKTKYNALKTDPKLOQAALP 179
QY 181 LVTVMQCROYWDSSTIDSMICAGGASCGSDSGGLVCQKGNVTWVLGI 231
Db 180 IVSEAKCKESWGSKITDVMICAGSGVSSCMGDSGGPLVCQKDGVTWLAGI 230
```

```
RESULT 9
Q9D8X8
ID Q9D8X8 PRELIMINARY; PRT; 263 AA.
AC Q9D8X8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 2200008D09RIK protein.
GN 2200008D09RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
```

```
SEQUENCE FROM N.A.
STRAIN=C57BL/6J; TISSUE=PANCREAS;
MEDLINE=21085660; PubMed=11217851;
Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
Blake J., Boffelli D., Bojunga N., Carninci P., Fujita M., de Bonaldo M.F.,
Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC EMBL: AK007566; BAB25112.1; -.
DR HSSP: P00766; IGCT.
DR MEROPS: S01.152; -.
DR MGD: MGI:1913723; 2200008D09RIK.
DR InterPro: IPR001314; Chymotrypsin.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PS02040; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 263 AA; 27898 MW; C0638FB8F905A92F CRC64;

Query Match 55.4%; Score 676.5; DB 11; Length 263;
Best Local Similarity 54.1%; Pred. No. 1.9e-53;
Matches 125; Conservative 37; Mismatches 68; Indels 1; Gaps 1;

QY 1 MLLSLTSLVLGGSGGCGIPALPSQRIYNGENAVLGSWPQVSLQDSSGFHFC 60
Db 1 MAFLWLVSFCALVGATFGCGVPAIQVLTGLSRIVNGEDAIPGSWPQVSLQDRTGFHFC 60
QY 61 GGSLSQSQSVVTAACHNVSFGRHFVVLGVEYDRSSNAEPLQVLSVSRATHPSWNSTTMN 120
Db 61 GGSLSSENWVTAACHGCVKT-TDVVVAGFDDQGSDEENVQVLKIAQVFNKFNSTFVRN 119
QY 121 DVTLLKASPAQYTRTRISPVCLASNEALTEGLTCVTTGWRGLSGVGNVTPAHLOQVALP 180
Db 120 DITLLKLATPAQFSETVSACVLTVDVDDPPATGLCATGTGKTKYNALKTDPKLOQAALP 179
QY 181 LVTVMQCROYWDSSTIDSMICAGGASCGSDSGGLVCQKGNVTWVLGI 231
Db 180 IVSEAKCKESWGSKITDVMICAGSGVSSCMGDSGGPLVCQKDGVTWLAGI 230

RESULT 10
Q9PWQ6
ID Q9PWQ6 PRELIMINARY; PRT; 263 AA.
AC Q9PWQ6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Chymotrypsin B precursor (EC 3.4.21.1).
GN CHYB.
OS Gadus morhua (Atlantic cod).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
 OX NCBI_TaxID=8049;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PYLORIC CAECA;
 RX MEDLINE=20464334; PubMed=11011764;
 RA Spilliaert R., Gudmundsdottir A.;
 RT "Molecular Cloning of the Atlantic Cod Chymotrypsinogen B.";
 RL Microb. Comp. Genomics 5:41-50(2000).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL; AJ242521; CAB43766.1; -.
 DR HSSP; P00766; 1CHG.
 DR MEROPS; S01.152; -.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Signal.
 FT SIGNAL 1 16 POTENTIAL.
 FT CHAIN 17 263 CHYMOTRYPSIN B.
 SQ SEQUENCE 263 AA; EF61B18A34EE5E7C CRC64;
 Query Match 55.2%; Score 674.5; DB 13; Length 263;
 Best Local Similarity 57.0%; Pred. No. 3e-53;
 Matches 127; Conservative 38; Mismatches 51; Indels 7; Gaps 5;
 QY 13 LGSSWGGGIPAIKPAIFSRIYNGENAVLGSPWQVSLQSSGFCGGLISQSWVVT 72
 DB 11 ISAIYGGGPAIQPVQYARIVNGEAVPHSPWQVSLQSSNGFCGGLINENWVVT 70
 QY 73 AAHCVSPGRHVFVVLGSDRS-NAEPLOVLSVSRATHPSWNSTTNNNDVTLKLKASPA 131
 DB 71 AAHCVNRT-YHRVIVGEHDKSRASDENIQILKPSMVTTPKWDSTINNDISLIKASPA 129
 QY 132 QYTRISPVCLASNEALTEGLCTVTTWG--RLSGVGNVTPAHLQVPLVTVNCRQ 189
 DB 130 VLGPNVSPVCLGESSDVFAPGMKCVTSCWGLTRYNAPG--TPNKLOQAALPLSMNECSQ 187
 QY 190 YW-DSSITSDMSICAGGASCGSDGGLVCGKGNWVLI 231
 DB 188 TWGNMISDVNICAGAGATSCWGDGGLVCGKGNWVLI 230
 RESULT 11
 Q9CQ52
 ID Q9CQ52 PRELIMINARY; PRT; 269 AA.
 AC Q9CQ52;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE 2310074F01R4K protein.
 GN ELA3B OR 2310074F01R4K.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=TONGUE;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Akakawa T., Shingagawa A., Shibata K., Konno H., Adachi J., Fukuda S.,
 RA Akakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Offelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL; AK010149; BAB26734.1; -.
 DR EMBL; AK009129; BAB26092.1; -.
 DR HSSP; P05805; IFON.
 DR MEROPS; S01.154; -.
 DR MGD; MGI:1915118; Ela3b.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 269 AA; 28904 MW; C543F76957B2A7CE CRC64;
 Query Match 38.7%; Score 472.5; DB 11; Length 269;
 Best Local Similarity 43.5%; Pred. No. 6.3e-35;
 Matches 104; Conservative 31; Mismatches 89; Indels 15; Gaps 7;
 QY 3 LLSLTLSVLVLLGSSWGGIPAIKPAIFSRIYNGENAVLGSPWQVSLQ---DSSGFHF 59
 DB 1 MLRLSSLLVALASGCGQSPHNP---SSRVNGEAVPHSPWQVSLQYKDGSPHHT 56
 QY 60 CGGSLISQSWVVTAAHCNVSPGRHVFVVLGSDRS-NAEPLOVLSVSR--ITHPNSNS-- 115
 DB 57 CGGSLITPDWVLVAGHCISTRTYQVVLGSHERGVEEQGVIPINAGDLFVHPKWNMSMC 116
 QY 116 TTNWNNVTLLKLKASPAQYTRISPVCLASNEALTEGLCTVTTWGRLSGVGNVTPAHLQ 175
 DB 117 VSGNDIALVKLSRSAGLDGAVQLCLPPAGEILPAGPCYISGWRGLSTNGPL-PDKLQ 175
 QY 176 QVALPLVTVNCRQ--YWDSSITDSMICAGGASCGSDGGLVCGKGN-TWVLI 231
 DB 176 QALLPVVDYEHCSRNNWGLSVKTTVMVACGDIQSGNGSDGGLNCPADNGTQVHV 234
 RESULT 12
 Q92077
 ID Q92077 PRELIMINARY; PRT; 266 AA.
 AC Q92077;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Elastase precursor (EC 3.4.21.37) (Leukocyte elastase) (Lysosomal
 DE elastase) (Neutrophil elastase) (Bone marrow serine protease)
 DE (MEDULLASIN).
 GN ELB.
 OS Gadus morhua (Atlantic cod).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
 OX NCBI_TaxID=8049;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96262679; PubMed=8925447;
 RA Gudmundsdottir E., Spilliaert R., Yang Q., Craik C.S., Bjarnason J.B.,
 RA Gudmundsdottir A.;

"Isolation and characterization of two cDNAs from Atlantic cod encoding two distinct psychrophilic elastases.";
Comp. Biochem. Physiol. 113B:795-801(1996).
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF PROTEINS, INCLUDING ELASTIN.
CC PREFERENTIAL CLEAVAGE: VAL-1-XAA > ALA-1-XAA.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY 1; ALSO KNOWN AS THE TRYPSIN FAMILY.
DR EMBL; U56956; BAB58350.1; -.
DR HSP; P05805; 1FON.
DR MEROPS; S01.155; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Signal; zymogen.
FT SIGNAL 1 15 POTENTIAL.
FT PROPEP 16 27 ACTIVATION PEPTIDE (POTENTIAL).
FT CHAIN 28 266 ELASTASE.
SQ SEQUENCE 266 AA; 28533 MW; B786B52C71559E2E CRC64;

Query Match 38.1%; Score 465.5; DB 13; Length 266;
Best Local Similarity 40.3%; Pred. No. 2.7e-34;
Matches 95; Conservative 45; Mismatches 83; Indels 13; Gaps 8;

Qy 4 LSLTSLVLGSSGCGIPAIKPAISQRTVINGENAVLGSPWQVSLQSSG---PHFC 60
Db 1 MKLLVFAFVAGAYGCGLPFPPIVT---RVYGGEDVRVHSPWQASLQYKSGNSFYHTC 57
Qy 61 GGSLSISQSVVTAACHNVSQGRHFVLCGEYD-RSSNAEPLQVLSVRAITHPNSNTTMN 119
Db 58 GGTLLIAQWVMTAAHC-IGSFYRVLLGKHNQDYNAGSLAIPAKTIIVHEKWDSSRIR 116
Qy 120 NDVTLKLPASPAQVTTTRISPVCLASSNEALTEGLTCVTGNGRLSGVGNVTPAHLQVVAL 179
Db 117 NDIALIKLASPDVSAITACVPDAEVLNLANGAPCVYTGNGRLWTGGPIADA-LQALL 175
Qy 180 PLVTNVNCRQY--WDSSITDSMICAGGAGA--SSCGDSSGGPLVCQKGN-TWVLIGI 231
Db 176 PVDAHCSRVDWNGSLVTTSMVCAAGDGVLASNGDSSGGLNCONADGSDVHGV 231

RESULT 13
Q9D7T9 PRELIMINARY; PRT; 269 AA.
ID Q9D7T9 PRELIMINARY; PRT; 269 AA.
AC Q9D7T9 PRELIMINARY; PRT; 269 AA.
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 2310074F01RIK protein
GN ELA3B OR 2310074F01RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=STOMACH;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Akizawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikolaio I., Pesole G., Quackenbush J.,
RA Schiml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz T., Whittaker C., Wilming L.,
RA Yushaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.
CC EMBL; AK008858; BAB25932.1; -.
DR HSP; P05805; 1FON.
DR MEROPS; S01.154; -.
DR MGD; MGI:1915118; Ela3b.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 269 AA; 28977 MW; 9F43F769DDB2A7CF CRC64;

Query Match 38.0%; Score 464.5; DB 11; Length 269;
Best Local Similarity 43.1%; Pred. No. 3.4e-34;
Matches 103; Conservative 31; Mismatches 90; Indels 15; Gaps 7;

Qy 3 LSLTSLVLGSSGCGIPAIKPAISQRTVINGENAVLGSPWQVSLQ---DSSGPHF 59
Db 1 MLRLSLLVALASGCEQSHNP---SSRVNCEAVPHSPWQVSLQYEKDGSEHHT 56
Qy 60 GGSLSISQSVVTAACHNVSQGRHFVLCGEYDNRSSNAEPLQVLSVRA--ITHPSWNS-- 115
Db 57 CGSLITPDVLTAGHCISTRTYOVVLGEHGEVEEQEQVPIPAGEILFVHPKWSMC 116
Qy 116 TTMNDVTLKLPASPAQVTTTRISPVCLASSNEALTEGLTCVTGNGRLSGVGNVTPAHLQ 175
Db 117 VSCNDIALVLSRSLQGLDAVOLACLPLPAGEILPAGPCYISGNGRLSTNGPL-PDKLQ 175
Qy 176 QVALPLVTNVNCRQ--YWDSSITDSMICAGGAGASCQDSSGGPLVCQKGN-TWVLIGI 231
Db 176 QALLPVVDYEHCSRWNWGLSVKTTMVCAGGDIQSGCNGDGGPLNCPADNQTQVHGV 234

RESULT 14
Q9W7P9 PRELIMINARY; PRT; 260 AA.
ID Q9W7P9 PRELIMINARY; PRT; 260 AA.
AC Q9W7P9 PRELIMINARY; PRT; 260 AA.
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Elastase 4 (Fragment).
OS Paralicthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidae; Paralicthidae; Paralicthys.
OX NCBI_TaxID=8255;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Suzuki T., Sivastava A.S., Kurokawa T.;
RT "Japanese flounder mRNA for elastase 4 precursor.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.
CC EMBL; AB029758; BAB2370.1; -.
DR HSP; P00761; 1EPT.
DR MEROPS; S01.155; -.
DR InterPro; IPR001314; Chymotrypsin.

DR InterPro: IPR001254; Ser_protease_Try.

DR Pfam: PF00089; trypsin; 1.

DR PRINTS: PR00722; CHYMOTRYPSIN.

DR SMART: SM00020; TRYP_SPC; 1.

DR PROSITE: PS00240; TRYPSIN_DOM; 1.

DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.

DR PROSITE: PS00135; TRYPSIN_SER; 1.

KW Hydrolase; Serine protease.

FT NON_TER 1

SQ SEQUENCE 260 AA; 27942 MW; 61883CD1B6A66075 CRC64;

Query Match 37.2%; Score 454.5; DB 13; Length 260;

Best Local Similarity 40.3%; Pred. No. 2.6e-33;

Matches 93; Conservative 45; Mismatches 80; Indels 13; Gaps 8;

QY 8 LSLVLGSSWGGIPALPALSFSQRIYNGENAVLGSMPQVSLQDSSG---FHFCGSL 64

DB 1 LALFVAG-AYGGLPTFFPTIT---RVGGDDVRENSPQVSLQKSGNFYTCGG 56

QY 65 ISQSWVTAAHCNVSPGRHFVVLGEYDRSSNAEPLQVLVSRAITHPSWNSTTMNNDVPL 124

DB 57 ISNQWLTAAHC-IGSRVRYVYMGKHLNDEAGSTAISPAKLIIVHNWDSYRINDIAL 115

QY 125 LKLASPAQYTTTRISPCVCLASSNEALTEGLTCVTGRLSGVGNVTPPAHLQOVALPLVTV 184

DB 116 IKLATPVTESDVMAACPLSDGIQPNGAPCVYTCGRLWTGPIADI-LQALLPVVDH 174

QY 185 NCRQ--YWDSSITSMICAGAG-ASSCGSGGPLVCQK-GNTWVLGI 231

DB 175 ATCTSDMWGSLVNSMVCAGGAGDLASCNGDSGGPLNCQSPDGSEVHG 225

RESULT 15

Q9DC82

ID Q9DC82 PRELIMINARY; PRT; 164 AA.

AC Q9DC82;

DT 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE 0910001G08Rik protein.

GN CTRL OR 0910001G08Rik.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=SPLEEN;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Salto R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,

RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

DR EMBL: AK003074; BAB22549.1;

DR HSSP: P00763; IDPO.

DR MEROPS: S01.997; -.

DR MGD: MGI:88558; Ctrl.

DR InterPro: IPR001254; Ser_protease_Try.

DR Pfam: PF00089; trypsin; 1.

DR SMART: SM00020; TRYP_SPC; 1.

DR PROSITE: PS00240; TRYPSIN_DOM; 1.

DR PROSITE: PS00135; TRYPSIN_SER; 1.

KW Hydrolase; Serine protease.

SQ SEQUENCE 164 AA; 17707 MW; 83791FD829ABEDD6 CRC64;

Query Match 36.6%; Score 447; DB 11; Length 164;

Best Local Similarity 69.4%; Pred. No. 7e-33;

Matches 86; Conservative 9; Mismatches 27; Indels 2; Gaps 2;

QY 110 HPSWNST-TMNDVTLKLIASP-AQYTTTRISPCVCLASSNEALTEGLTCVTGWRLSGVG 167

DB 8 HPHWGGPHPLKTRLFLEFFFPPEKTKFPFGLASTNEALPSGLTCVTGWRISGVG 67

QY 168 NYTPAHLQOVALPLTVNOCROYWDSSTIDSMICAGGAGSSCGSGGPLVCQKNTWV 227

DB 68 NYTPARLQOVLPLTVNOCROYWGARIITDAMICAGGSGASSCGSGGPLVCQKNTWV 127

QY 228 LIGI 231

DB 128 LIGI 131

Search completed: December 20, 2002, 15:18:53

Job time : 29.6087 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:12:58 ; Search time 7.91304 Seconds
(without alignments)
1210.790 Million cell updates/sec

Title: US-09-856-319B-2_COPY_1_231

Perfect score: 1221

Sequence: 1 MLLSLTSLVLLGSGWCG.....GDSGGPLVCQKGTWVLIGI 231

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1214	99.4	264	1	CTRL_HUMAN
2	714.5	58.5	263	1	CTR2_CANFA
3	701.5	57.5	263	1	CTR2_GADMO
4	689.5	56.5	263	1	CTR3_HUMAN
5	670.5	54.9	263	1	CTR3_RAT
6	666.5	54.6	245	1	CTR3_BOVIN
7	684.5	54.4	245	1	CTR3_GADMO
8	657.5	53.8	245	1	CTR3_BOVIN
9	472	38.7	269	1	EL2_PIG
10	453.5	37.1	270	1	EL3B_HUMAN
11	453	37.1	269	1	EL2_BOVIN
12	445	36.4	269	1	EL2A_HUMAN
13	438	35.9	271	1	EL2_RAT
14	427	35.0	269	1	EL2B_HUMAN
15	420.5	34.4	270	1	EL3A_HUMAN
16	415	34.0	271	1	EL2_MOUSE
17	410.5	33.6	253	1	CAC3_BOVIN
18	410.5	33.6	343	1	PSS8_HUMAN
19	408	33.4	454	1	TMS2_HUMAN
20	407.5	33.4	268	1	CLCR_RAT
21	404	33.1	343	1	PLMN_SHEEP
22	398.5	32.6	342	1	PSS8_RAT
23	397	32.5	810	1	PLMN_HUMAN
24	396.5	32.5	342	1	PSS8_MOUSE
25	396	32.4	810	1	PLMN_MACMU
26	395.5	32.4	1035	1	ENTK_BOVIN
27	392	32.1	266	1	EL1_BOVIN
28	391	32.0	790	1	PLMN_PIG
29	390.5	32.0	268	1	CLCR_HUMAN
30	389.5	31.9	812	1	PLMN_MOUSE
31	389	31.9	338	1	PLMN_HORSE
32	388	31.8	270	1	TRY1_MERON
33	386	31.6	786	1	STUB_DROME

34	384.5	31.5	238	1	TRY3_SALSA
35	384.5	31.5	273	1	MCT7_MOUSE
36	383.5	31.4	246	1	TRY1_CANFA
37	383.5	31.4	246	1	TRY1_RAT
38	383.5	31.4	275	1	TRY1_PIG
39	382.5	31.3	1069	1	ENTK_MOUSE
40	382	31.3	248	1	TRY1_CHICK
41	382	31.3	248	1	TRY2_CHICK
42	382	31.3	1019	1	ENTK_HUMAN
43	380	31.1	333	1	PLMN_CANFA
44	379	31.0	248	1	TRY3_CHICK
45	379	31.0	492	1	TMS2_HUMAN

P35033 salmo salar
Q02844 mus musculus
P06871 canis fam1
P00762 rattus norv
Q9n2d1 sus scrofa
P97435 mus musculus
Q90628 gallus gall
Q90627 gallus gall
P98073 homo sapien
P80009 canis fam1
Q90629 gallus gall
O15393 homo sapien

ALIGNMENTS

RESULT 1	CTRL_HUMAN	STANDARD;	PRT;	264 AA.
ID	CTRL_HUMAN			
AC	P40313			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Chymotrypsin-like protease CTRL-1 precursor (EC 3.4.21.-)			
GN	CTRL OR CTRL1			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94093544; PubMed=8268911;			
RA	Larsen F., Solheim J., Kristensen T., Kolsto A.B., Prydz H.;			
RT	"A tight cluster of five unrelated human genes on chromosome			
RT	16q22.1.";			
RL	Hum. Mol. Genet. 2:1589-1595(1993).			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X71874; CAA50710.1; -			
DR	EMBL; X71877; CAA50711.1; -			
DR	HSSP; P00763; LDPO.			
DR	MEROPS; S01-256; -			
DR	Genew; HGNC:2524; CTRL.			
DR	MIM; 118888; -			
DR	InterPro; IPR001314; Chymotrypsin.			
DR	InterPro; IPR001254; Ser_protease_Try.			
DR	Pfam; PF00089; trypsin; 1.			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	SMART; SM00020; Tryp_Spc; 1.			
DR	PROSITE; PS50240; TRYPsin_DOM; 1.			
DR	PROSITE; PS00134; TRYPsin_HIS; 1.			
DR	PROSITE; PS00135; TRYPsin_SER; 1.			
KW	Hydrolase; Serine protease; Glycoprotein; zymogen; Signal.			
FT	SIGNAL 1 18			
FT	PROPEP 19 33			
FT	CHAIN 34 264			
FT	ACT_SITE 75 75			
FT	ACT_SITE 121 121			
FT	ACT_SITE 214 214			
FT	CARBOHYD 114 114			
FT	DISULFID 19 141			
FT	DISULFID 60 76			
FT	DISULFID 155 220			
FT	DISULFID 187 201			

POTENTIAL.
ACTIVATION PEPTIDE (POTENTIAL).
CHYMOTRYPSIN-LIKE PROTEASE CTRL-1.
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
N-LINKED (GLCNAC. . .) (POTENTIAL).
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.

FT	CHAIN	34	164	CHYMOTRYPSIN 2, B CHAIN.
FT	CHAIN	167	263	CHYMOTRYPSIN 2, C CHAIN.
FT	ACT_SITE	75	75	CHARGE RELAY SYSTEM.
FT	ACT_SITE	120	120	CHARGE RELAY SYSTEM.
FT	ACT_SITE	213	213	CHARGE RELAY SYSTEM.
FT	DISULFID	19	140	BY SIMILARITY.
FT	DISULFID	60	76	BY SIMILARITY.
FT	DISULFID	154	219	BY SIMILARITY.
FT	DISULFID	186	200	BY SIMILARITY.
FT	DISULFID	209	238	BY SIMILARITY.
SEQ	SEQUENCE	263 AA;	27787 MW;	2A2F449D813B3961 CRC64;

Query Match 58.5%; Score 714.5; DB 1; Length 263;
Best Local Similarity 56.7%; Pred. No. 9.7e-55;
Matches 131; Conservative 35; Mismatches 64; Indels 1; Gaps

QY	1	MLLSLTSLVLGSSGGCGIPAKIPALSPQRSIVGNNALVGSWPQVSLDSSGSHFC	60
QY <td>1</td> <td>MLLSLTSLVLGSSGGCGIPAKIPALSPQRSIVGNNALVGSWPQVSLDSSGSHFC</td> <td>60</td>	1	MLLSLTSLVLGSSGGCGIPAKIPALSPQRSIVGNNALVGSWPQVSLDSSGSHFC	60
DB	1	MAFWLLSCFALIGTAFGCVPAIQVLSGLSRIVGEDAVPGSWPQVSLDSSGSHFC	60
QY <td>61</td> <td>GGSLISDQVVTAAHCNVSPGRHFVVLGEXDRSSNAEPQLVLSVSRATHPSWNSTMMN</td> <td>120</td>	61	GGSLISDQVVTAAHCNVSPGRHFVVLGEXDRSSNAEPQLVLSVSRATHPSWNSTMMN	120
DB	61	GGSLISDQVVTAAHCNVSPGRHFVVLGEXDRSSNAEPQLVLSVSRATHPSWNSTMMN	120
QY <td>121</td> <td>DVTLLKLASPAQYTRISPCVLASSNEALTEGUTCTVTTGWRGLSVGVNTPAHLQOVALP</td> <td>180</td>	121	DVTLLKLASPAQYTRISPCVLASSNEALTEGUTCTVTTGWRGLSVGVNTPAHLQOVALP	180
DB	120	DITLLKLATARESKTVSAVCLPQATDDFPAGYLCVTTGWLTKHTNANTPKLQQAALP	179
QY <td>181</td> <td>LVTVNQCRQYWDSSITDSMTICAGGAGSSCGDGGPLVCQKGNWVLGI</td> <td>231</td>	181	LVTVNQCRQYWDSSITDSMTICAGGAGSSCGDGGPLVCQKGNWVLGI	231
DB <td>180</td> <td>LLSNAECKFGWSKITDLWCAGAGSVSSCMGDSGGPLVCQKDGAWTLVGI</td> <td>230</td>	180	LLSNAECKFGWSKITDLWCAGAGSVSSCMGDSGGPLVCQKDGAWTLVGI	230

RESULT 3

ID	CTRA_GADMO	STANDARD;	PRT;	263 AA.
AC	P47796;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DE	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Chymotrypsin A precursor (EC 3.4.21.1).			
OS	Gadus morhua (Atlantic cod).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;			
OC	Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.			
OX	NCBI_TaxID=8049;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Pyloric caeca;			
RX	MEDLINE=94368860; PubMed=8086467;			
RA	Gudmundsdottir A., Oskarsson S., Eakin A.E., Craik C.S.,			
RA	Bjarnason J.B.;			
RT	"Atlantic cod cDNA encoding a psychrophilic chymotrypsinogen.";			
RL	Biochim. Biophys. Acta 1219:211-214(1994).			
RN	[2]			
RP	SEQUENCE OF 19-30 AND 34-49.			
RC	TISSUE=Pyloric caeca;			
RX	MEDLINE=92111252; PubMed=1764912;			
RA	Ageirsson B., Bjarnason J.B.;			
RT	"Structural and kinetic properties of chymotrypsin from Atlantic cod			
RT	(Gadus morhua). Comparison with bovine chymotrypsin.";			
RL	Comp. Biochem. Physiol. 99B:327-335(1991).			
CC	-1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,			
CC	Phe-I-Xaa, Leu-I-Xaa.			
CC	-1- SUBCELLULAR LOCATION: Extracellular.			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			

This SWISS-PROT entry is copyright. It is produced through a collabor
between the Swiss Institute of Bioinformatics and the EMBL outstat
the European Bioinformatics Institute. There are no restrictions on
use by non-profit institutions as long as its content is in no
modified and this statement is not removed. Usage by and for commu
entities requires a license agreement (See <http://www.isb-sib.ch/anno>)

or send an email to license@isb-sib.ch).

```
CC EMBL; X78490; CAA55242.1; -
CC HSSP; P00766; ICHG.
DR MEROPS; S01.152; -.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYP_SPC; 1.
DR PROSITE; PS00134; TRYP_SIN_HIS; 1.
DR PROSITE; PS00135; TRYP_SIN_SER; 1.
KW Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
FT SIGNAL 1 18
FT CHAIN 19 263
FT ACT_SITE 75 75 CHYMOTRYPSIN A.
FT ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 19 140 BY SIMILARITY.
FT DISULFID 60 76 BY SIMILARITY.
FT DISULFID 154 219 BY SIMILARITY.
FT DISULFID 186 200 BY SIMILARITY.
FT DISULFID 209 238 BY SIMILARITY.
FT CONFLICT 21 21 R -> S (IN REF. 2).
FT CONFLICT 25 25 S -> Q (IN REF. 2).
FT CONFLICT 29 29 T -> S (IN REF. 2).
FT CONFLICT 44 44 S -> T (IN REF. 2).
FT CONFLICT 46 46 S -> Y (IN REF. 2).
SQ SEQUENCE 263 AA; 28294 MW; 47AAC699A0A64FBB CRC64;

Query Match 57.5%; Score 701.5; DB 1; Length 263;
Best Local Similarity 58.6%; Pred. No. 1.3e-53;
Matches 129; Conservative 33; Mismatches 57; Indels 1; Gaps 1;

Qy 12 LLGSSWGGIPAIKPAISFQRIYNGENAVLGSPWQVSLQDSSGFHFCGGLISQSWV 71
Db 12 LFRRTYGGCPAIPVITGYSRIYNGEAVPHSWQVSLQDQGTGFHFCGGLINENWV 71
Qy 72 TAAHCNVSPGRHVVGLGYDRSSNAEPLOVLSVSRATHPSWNSNTMNDVTLKLASPA 131
Db 72 TAAHCNVK-NYHRVVLGDEHSSSEGQVMTVGQVFKPRYNGFTINDILLVLAATPA 130
Qy 132 QYTRISPVCLASNEALTEGLCTVTTGWRGLSGVGNVTPAHQQVALPLVTVNQCRYW 191
Db 131 TLNMRVSPVCLAEETDDVPEGGMKCVTSGMGLTRYNAADTPALQQAALPLLTNECKKEW 190
Qy 192 DSSITDSMICAGGACSCQDSSGGPLVCQKGNWTLVGI 231
Db 191 GNKISDLMICAGAGACSCMGDSGGPLVCQKAGSWTLVGI 230
```

RESULT 4

CTRB_HUMAN STANDARD; PRT; 263 AA.

```
AC P17538;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsinogen B precursor (EC 3.4.21.1).
GN CTRB1 OR CTRB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Pancreas;
RX MEDLINE=89134264; PubMed=2917002;
RA Tomita N., Izumoto Y., Horii A., Doi S., Yokouchi H., Ogawa M.,
RA Mori T., Matsubara K.;
RT "Molecular cloning and nucleotide sequence of human pancreatic
chymotrypsinogen cDNA.";
```

```
RL Biochem. Biophys. Res. Commun. 158:569-575(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Pancreas;
RA Strausberg R.;
RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
CC Phe-I-Xaa, Leu-I-Xaa;
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M24400; AAA52128.1; -.
CC EMBL; BC005385; AAH05385.1; -.
CC PIR; A31299; A31299.
CC HSSP; P00766; ICHG.
CC MEROPS; S01.152; -.
CC MIN; I18890; -.
CC InterPro; IPR001314; Chymotrypsin.
CC IncerPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC SMART; SM00020; Tryp_SPC; 1.
CC PROSITE; PS00240; TRYP_SIN_DOM; 1.
CC PROSITE; PS00134; TRYP_SIN_HIS; 1.
CC PROSITE; PS00135; TRYP_SIN_SER; 1.
KW Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
FT SIGNAL 1 18
FT CHAIN 19 263 CHYMOTRYPSINOGEN B.
FT CHAIN 19 31 CHYMOTRYPSIN B, A CHAIN.
FT CHAIN 34 164 CHYMOTRYPSIN B, B CHAIN.
FT CHAIN 167 263 CHYMOTRYPSIN B, C CHAIN.
FT ACT_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 19 140 BY SIMILARITY.
FT DISULFID 60 76 BY SIMILARITY.
FT DISULFID 154 219 BY SIMILARITY.
FT DISULFID 186 200 BY SIMILARITY.
FT DISULFID 209 238 BY SIMILARITY.
SQ SEQUENCE 263 AA; 27870 MW; 4C1C055A490B8701 CRC64;

Query Match 56.5%; Score 689.5; DB 1; Length 263;
Best Local Similarity 55.4%; Pred. No. 1.4e-52;
Matches 128; Conservative 35; Mismatches 67; Indels 1; Gaps 1;

Qy 1 MLLSLTSLVLLGSSWGGIPAIKPAISFQRIYNGENAVLGSPWQVSLQDSSGFHFC 60
Db 1 MAFLWLLSCWALLGTTFGCGVPAIHPVLVSLGSLRVNGEDAVPGSPWQVSLQDQGTGFHFC 60
Qy 61 GGSLSISQSWVYTAHCNVSPGRHVVGLGYDRSSNAEPLOVLSVSRATHPSWNSNTMNN 120
Db 61 GGSLSIEDWVYTAHCHGVRTS-DVYVAGEFDQGSDEENIQVLKIAKVPKPKFSILTWN 119
Qy 121 DVTLLKASPAGYTRISPVCLASNEALTEGLCTVTTGWRGLSGVGNVTPAHQQVALP 180
Db 120 DITLLKATPARFSQTVSAVCLPSADDDPPAGTLCATGWTGKTKYKANKTPDKLQQAALP 179
Qy 181 LVTNVQCRYWDSSTDSMICAGGACSCQDSSGGPLVCQKGNWTLVGI 231
Db 180 LLSNAECKKSGWRRITDVMICAGAGVSCMGDSGGPLVCQKDGAWTLVGI 230

RESULT 5
CTRB_RAT
```

```
ID CTRB_RAT STANDARD; PRT; 263 AA.
AC P07338;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsinogen B precursor (EC 3.4.21.1).
GN CTRB1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85054881; PubMed=6209274;
RA Bell G.I., Quinto C., Quiroga M., Valenzuela P., Craik C.S.,
RT Rutter W.J.;
RT "Isolation and sequence of a rat chymotrypsin B gene.";
RL J. Biol. Chem. 259:14265-14270(1984).
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
phe-I-Xaa, Leu-I-Xaa.
CC -|- SUBCELLULAR LOCATION: Extracellular.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02298; AAA98732.1; -.
DR PIR; A22658; KYRUB.
DR HSP; P00766; 1CHG.
DR MEROPS; S01.152; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
FT SIGNAL 1 18
FT CHAIN 19 263 CHYMOTRYPSINOGEN B.
FT CHAIN 19 31 CHYMOTRYPSIN B, A CHAIN.
FT CHAIN 34 164 CHYMOTRYPSIN B, B CHAIN.
FT CHAIN 167 263 CHYMOTRYPSIN B, C CHAIN.
FT ACT_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 19 140 BY SIMILARITY.
FT DISULFID 60 76 BY SIMILARITY.
FT DISULFID 154 219 BY SIMILARITY.
FT DISULFID 186 200 BY SIMILARITY.
FT DISULFID 209 238 BY SIMILARITY.
SQ SEQUENCE 263 AA; 27849 MW; ACAFDACF8C4DA6D CRC64;

Query Match 54.9%; Score 670.5; DB 1; Length 263;
Best Local Similarity 53.2%; Pred. No. 6.1e-51;
Matches 123; Conservative 35; Mismatches 72; Indels 1; Gaps 1;

QY 1 MLLSLTLVLGGSGWGCGIPAIKPAISFQRIYNGENAVLGSWPQVSLQDSSGFHFC 60
DB 1 MAFLWLVSCFALVATFGCGVTIQVTLGLSRVNGEDAIPCSWPQVSLQDKTGFHFC 60
QY 61 GGSLSQSQVWVTAACHNVPGRHFVYLGEYDRSSNAEPLQVLSVSRATHPSWNNSTMMN 120
DB 61 GGSLSQVWVTAACHNVPGRHFVYLGEYDRSSNAEPLQVLSVSRATHPSWNNSTMMN 119
QY 121 DVTLLKASPAQYTTTRISPYCLASSNEALTEGLTCVTTGWLGSVGNVTPAHQLQVALP 180
DB 121 DVTLLKASPAQYTTTRISPYCLASSNEALTEGLTCVTTGWLGSVGNVTPAHQLQVALP 180
```

```
DB 120 DITLLKATPAQFSFVSACVLPNVDDPPPTGTCATGKTKYNALKTPKLOQAALP 179
QY 181 LVTVNQCROYWDSITDSMICAGGASSCGQSGSGLVLCQKGTWVLGI 231
DB 180 IVSEADCKKSGSKITDYMTWCAGSSVSCMGDSGGLVLCQKQDGVWTLGI 230

RESULT 6
ID CTRB_BOVIN STANDARD; PRT; 245 AA.
AC P00767;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE Chymotrypsinogen B (EC 3.4.21.1).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE, DISULFIDE BONDS, AND ACTIVE SITE.
RX MEDLINE=68238908; PubMed=5649671;
RA Smillie L.B., Furka A., Nagabhushan N., Stevenson K.J., Parkes C.O.;
RT "Structure of chymotrypsinogen B compared with chymotrypsinogen A and
trypsinogen.";
RL Nature 218:343-346(1968).
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
phe-I-Xaa, Leu-I-Xaa.
CC -|- SUBCELLULAR LOCATION: Extracellular.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -|- DATABASE: NAME=Washington enzyme manual;
CC WWW="http://www.worthington-biochem.com/manual/C/CHY.html".
DR PIR; A00953; KYBOB.
DR HSP; P00766; 1ACB.
DR MEROPS; S01.152; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen.
FT CHAIN 1 13 CHYMOTRYPSIN B, A CHAIN.
FT CHAIN 16 146 CHYMOTRYPSIN B, B CHAIN.
FT CHAIN 149 245 CHYMOTRYPSIN B, C CHAIN.
FT ACT_SITE 57 57 CHARGE RELAY SYSTEM.
FT ACT_SITE 102 102 CHARGE RELAY SYSTEM.
FT ACT_SITE 195 195 CHARGE RELAY SYSTEM.
FT DISULFID 1 122 CHARGE RELAY SYSTEM.
FT DISULFID 42 58
FT DISULFID 136 201
FT DISULFID 168 182
FT DISULFID 191 220
SQ SEQUENCE 245 AA; 25755 MW; 678016446FF5FEB5 CRC64;

Query Match 54.6%; Score 666.5; DB 1; Length 245;
Best Local Similarity 57.3%; Pred. No. 1.3e-50;
Matches 122; Conservative 28; Mismatches 62; Indels 1; Gaps 1;

QY 19 CGIPAIPALSPQSRIYNGENAVLGSWPQVSLQDSSGFHFCGSLISQSWVTAACHNV 78
DB 1 CGVPAIQPVLSGLRVNGEDAVPGSWPQVSLQDSTGFHFCGSLISQSWVTAACHNV 60
QY 79 SGRHFVVLGEYDRSSNAEPLQVLSVSRATHPSWNNSTMMNNDVTLKASPAQYTTTRIS 138
DB 61 TTS-DVVVAGEFDQGLETDQTLKIGKVPKPKFKPESILTVRNDITLLKATPAQFSFVS 119
QY 139 PYCLASSNEALTEGLTCVTTGWLGSVGNVTPAHQLQVALPVTNOCROYWDSITDS 198
DB 120 AVCLPSADEDFPAGMLCATTGKTKYNALKTPDKLQQAATLPIVSTDCRKYWGSRTDV 179
```

```

QY 199 MICAGGAGSSCGDGGPLVCQKGNWTWVLGI 231
DB 180 MICAGASGVSSCMGDSGGPLVCQKGNWTWVLGI 212

RESULT 7
CTRB_GADMO          STANDARD;          PRT;    245 AA.
ID CTRB_GADMO
AC P80646;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsin B (EC 3.4.21.1).
OS Gadus morhua (Atlantic cod).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Neoteleostei;
OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
OX NCBI_TaxID=8049;
RN [1]
RP SEQUENCE.
RC TISSUE=Pyloric caeca;
RX MEDLINE=96439045; PubMed=8841380;
RA Leth-Larsen R., Asgeirsson B., Thorolfsson M., Noerregaard-Madsen M.,
RA Hoejrup P.;
RT "Structure of chymotrypsin variant B from Atlantic cod, Gadus
RT morhua."
RL Biochim. Biophys. Acta 1297:49-56(1996).
RN [2]
RP SEQUENCE OF 1-12 AND 16-31.
RC TISSUE=Pyloric caeca;
RX MEDLINE=92111252; PubMed=1764912;
RA Asgeirsson B., Bjarnason J.B.;
RT "Structural and kinetic properties of chymotrypsin from Atlantic cod
RT (Gadus morhua). Comparison with bovine chymotrypsin."
RL Comp. Biochem. Physiol. 99B:327-335(1991).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
CC Phe-I-Xaa, Leu-I-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR HSP: P00766; 1CHG.
DR MEROPS: S01.152; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PS02040; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen.
FT CHAIN 1 13 CHYMOTRYPSIN B, A CHAIN.
FT CHAIN 16 245 CHYMOTRYPSIN B, B CHAIN.
FT ACT_SITE 57 57 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 101 101 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 195 195 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 1 121 BY SIMILARITY.
FT DISULFID 42 58 BY SIMILARITY.
FT DISULFID 135 201 BY SIMILARITY.
FT DISULFID 167 182 BY SIMILARITY.
FT DISULFID 191 220 BY SIMILARITY.
FT CONFLICT 9 11 QVT -> VIS (IN REF. 2).
FT CONFLICT 26 26 S -> T (IN REF. 2).
FT CONFLICT 28 29 PW -> Y (IN REF. 2).
SQ SEQUENCE 245 AA; 26260 MW; 74FE0D425517AB02 CRC64;

Query Match 54.4%; Score 664.5; DB 1; Length 245;
Best Local Similarity 58.3%; Pred. No. 1.9e-50;
Matches 126; Conservative 34; Mismatches 49; Indels 7; Gaps 5;

QY 19 CGIPAIKPAISQRIYNGENAVLSWQVSLQSSGFHFCGGLISGSWVYVTAHCNV 78
DB 1 CGSPAIOQVGTGYRIVNGEAVPHSWQVSLQSSGFHFCGGLISGSWVYVTAHCNV 60

```

```

QY 79 SPGRFVVLGEYDRSSNAEPLOVLSVSRATHTPSWNSTTMNDVTLTKLASPAQVYTRIS 138
DB 61 RT-YHRVIVGEHDKASD-ENIQILKPSVFTHPKWDSTRIINNDISLIKLASPAVLGTNVS 118

QY 139 PVLCLASSNEALTEGLTCVTTGNG--RLSGVGNVTPAHLQOVALPLVTVNOCROYW-DSIS 195
DB 119 PVLCELLSSDVFAPGMKCVTSGMGLTRYNAPG--TPNKLQQAALPLMSNEECSTQWGNMI 176

QY 196 TDSMICAGAGASSCGDGGPLVCQKGNWTWVLGI 231
DB 177 SDVMICAGAAGATSCMGDSGGPLVCQKGNWTWVLGI 212

RESULT 8
CTRB_BOVIN          STANDARD;          PRT;    245 AA.
ID CTRB_BOVIN
AC P00766;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsinogen A (EC 3.4.21.1).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE, DISULFIDE BONDS, AND ACTIVE SITE.
RX MEDLINE=67181721; PubMed=5971783;
RA Brown J.R., Hartley B.S.;
RT "Location of disulphide bridges by diagonal paper electrophoresis.
RT The disulphide bridges of bovine chymotrypsinogen A."
RL Biochem. J. 101:214-228(1966).
RN [2]
RP REVISION TO 102.
RX MEDLINE=69106266; PubMed=5764436;
RA Blow D.M., Birktoft J.J., Hartley B.S.;
RT "Role of a buried acid group in the mechanism of action of
RT chymotrypsin."
RL Nature 221:337-340(1969).
RN [3]
RP PRELIMINARY SEQUENCE.
RA Hartley B.S.;
RT "Amino-acid sequence of bovine chymotrypsinogen-A."
RL Nature 201:1284-1287(1964).
RN [4]
RP SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=67183948; PubMed=5972866;
RA Meloun B., Klueh I., Kostka V., Moravsek L., Prusik Z., Vanacek J.,
RA Kell B., Sorm F.;
RT "Covalent structure of bovine chymotrypsinogen A."
RL Biochim. Biophys. Acta 130:543-546(1966).
RN [5]
RP ACTIVE SITE.
RX MEDLINE=67181723; PubMed=5971785;
RA Smillie L.B., Hartley B.S.;
RT "Histidine sequences in the active centres of some 'serine'
RT proteinases."
RL Biochem. J. 101:232-241(1966).
RN [6]
RP X-RAY CRYSTALLOGRAPHY.
RX MEDLINE=7203052; PubMed=4399050;
RA Birktoft J.J., Blow D.M., Henderson R., Steitz T.A.;
RT "I. Serine Proteinases. The structure of alpha-chymotrypsin."
RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 257:67-76(1970).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF CHYMOTRYPSINOGEN.
RX MEDLINE=7017557; PubMed=5442169;
RA Freer S.T., Kraut J., Robertus J.D., Wright H.T., Xuong N.H.;
RT "Chymotrypsinogen: 2.5-A crystal structure, comparison with alpha-
RT chymotrypsin, and implications for zymogen activation."
RL Biochemistry 9:1997-2009(1970).

```

RN	[8]	X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF GAMMA-CHYMOTRYPSIN.
RX	MDLNAME=82078042; PubMed=6914398;	
RC	Cohen G.H., Silvertown E.W., Davies D.R.;	
RT	"Refined crystal structure of gamma-chymotrypsin at 1.9-A resolution.	
RT	Comparison with other pancreatic serine proteases.";	
RL	J. Mol. Biol. 148:449-479(1981).	
CC	[9]	
RN	X-RAY CRYSTALLOGRAPHY (1.68 ANGSTROMS) OF ALPHA-CHYMOTRYPSIN.	
RX	MDLNAME=86011575; PubMed=4046030;	
RC	Tsakuda H., Blow D.M.;	
RT	"Structure of alpha-chymotrypsin refined at 1.68-A resolution."	
RL	J. Mol. Biol. 184:703-711(1985)..	
CC	- - CATALYTIC ACTIVITY: Preferential cleavage: Tyr- -Xaa, Trp- -Xaa,	
CC	Phe- -Xaa, Leu- -Xaa.	
CC	- - SUBCELLULAR LOCATION: Extracellular.	
CC	- - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.	
CC	- - DATABASE: NAME=Worthington enzyme manual;	
CC	WWW="http://www.worthington-biochem.com/manual/C/CHY.html".	
DR	PTR; A00952; KBOA.	
DR	PDB; 2CGA; 15-APR-90.	
DR	PDB; 2CHA; 31-MAY-84.	
DR	PDB; 4CHA; 29-OCT-85.	
DR	PDB; 5CHA; 16-OCT-87.	
DR	PDB; 6CHA; 16-OCT-87.	
DR	PDB; 1CHG; 27-JAN-84.	
DR	PDB; 1CHO; 16-JUL-88.	
DR	PDB; 2GCH; 31-MAY-84.	
DR	PDB; 3GCH; 15-OCT-92.	
DR	PDB; 4GCH; 15-OCT-90.	
DR	PDB; 5GCH; 15-OCT-90.	
DR	PDB; 6GCH; 15-OCT-90.	
DR	PDB; 7GCH; 15-OCT-90.	
DR	PDB; 8GCH; 15-JUL-93.	
DR	PDB; 1GCT; 15-OCT-91.	
DR	PDB; 2GCT; 15-OCT-91.	
DR	PDB; 3GCT; 15-OCT-91.	
DR	PDB; 1ACB; 31-OCT-93.	
DR	PDB; 1GMC; 31-OCT-93.	
DR	PDB; 1GMD; 31-OCT-93.	
DR	PDB; 1CGI; 30-APR-94.	
DR	PDB; 1CGJ; 30-APR-94.	
DR	PDB; 1GCD; 22-JUN-94.	
DR	PDB; 1GHA; 22-JUN-94.	
DR	PDB; 1GHB; 22-JUN-94.	
DR	PDB; 1GMH; 30-SEP-94.	
DR	PDB; 2GMT; 01-NOV-94.	
DR	PDB; 1MTN; 17-AUG-96.	
DR	PDB; 1AB9; 20-AUG-97.	
DR	PDB; 1AFQ; 17-SEP-97.	
DR	PDB; 1CAO; 23-JUL-97.	
DR	PDB; 1CBW; 23-JUL-97.	
DR	PDB; 1VGC; 12-NOV-97.	
DR	PDB; 2VGC; 12-NOV-97.	
DR	PDB; 3VGC; 12-NOV-97.	
DR	PDB; 4VGC; 12-NOV-97.	
DR	PDB; 1HJA; 14-JAN-98.	
DR	MEROPS; S01.001; -.	
DR	InterPro; IPR001314; Chymotrypsin.	
DR	InterPro; IPR001254; Ser.protease_Try.	
DR	Pfam; PF000089; trypsin; 1.	
DR	PRINTS; PR00722; CHYMOTRYPSIN.	
DR	SMART; SMO0020; Tryp_Spc; 1.	
DR	PROSITE; PS0240; TRYPSIN_DOM; 1.	
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.	
DR	PROSITE; PS00135; TRYPSIN_SER; 1.	
KW	Hydrolase; Serine protease; Digestion; Pancreas; Zymogen;	
KW	3D-structure.	
FT	CHAIN 1 13 CHYMOTRYPSIN A, A CHAIN.	
FT	CHAIN 16 146 CHYMOTRYPSIN A, B CHAIN.	
FT	CHAIN 149 245 CHYMOTRYPSIN A, C CHAIN.	
FT	ACT_SITE 57 57 CHARGE RELAY SYSTEM.	
FT	ACT_SITE 102 102 CHARGE RELAY SYSTEM.	

ID	EL2_PIG	STANDARD;	PRT;	269 AA.
AC	P08419;			
DT	01-AUG-1988 (Rel. 08, Created)			
DT	01-AUG-1988 (Rel. 08, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DT	Elastase 2 precursor (EC 3.4.21.71).			
DE	ELA2.			
GN	Elus scrofa (Pig).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
ON	NCBI_TaxID=9823;			
OX	[1]			
RP	SEQUENCE FROM N.A.			
RR	MEDLINE=87217962; PubMed=3646943;			
RX	Kawashima I., Tani T., Shimoda K., Takiguchi Y.;			
RT	"Characterization of pancreatic elastase II cDNAs: two elastase II			
RT	mRNAs are expressed in human pancreas.";			
RL	DNA 6:163-172(1987).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88198076; PubMed=2834346;			
RX	Shirasu Y., Yoshida H., Matsuki S., Takemura K., Ikeda N.,			
RA	Shimada Y., Ozawa T., Mikayama T., Iijima H., Ishida A., Sato Y.,			
RA	Tamai Y., Tanaka J., Ikenaga H.;			
RT	"Molecular cloning and expression in Escherichia coli of a cDNA			
RT	encoding human pancreatic elastase 2.";			
RL	J. Biochem. 102:1555-1563(1987).			
CC	-1- FUNCTION: ACTS UPON ELASTIN.			
CC	-1- CATALYTIC ACTIVITY: Preferential cleavage: Leu- -Xaa, Met- -Xaa			
CC	and Phe- -Xaa. Hydrolyzes elastin.			
CC	-1- SUBCELLULAR LOCATION: Secreted.			
CC	-1- TISSUE SPECIFICITY: PANCREAS.			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@lsb-sib.ch).			
CC	-----			
DR	EMBL; M16651; AAA31027.1; -;			
DR	EMBL; D00237; BAA00166.1; -;			
DR	PIR; A26823; A26823.			
DR	HSSP; P00763; LDPO.			
DR	MEROPS: S01.155; -;			
DR	InterPro: IPR001314; Chymotrypsin.			
DR	InterPro: IPR001254; Ser_protease_Try.			
DR	Pfam; PF00089; trypsin; 1.			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	SMART; SM00020; TRYPSIN; 1.			
DR	PROSITE; PS02040; TRYPSIN_DOM; 1.			
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.			
DR	PROSITE; PS00135; TRYPSIN_SER; 1.			
KW	Hydrolase; Serine protease; zymogen; Signal.			
FT	SIGNAL	1	16	
FT	PROPEP	17	28	ACTIVATION PEPTIDE.
FT	CHAIN	29	269	ELASTASE 2.
FT	ACT_SITE	73	73	
FT	ACT_SITE	121	121	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	216	216	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	DISULFID	58	74	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	DISULFID	155	222	BY SIMILARITY.
FT	DISULFID	186	202	BY SIMILARITY.
FT	DISULFID	212	243	BY SIMILARITY.
FT	CONFLICT	10	10	BY SIMILARITY.
FT	CONFLICT	118	118	L -> S (IN REF. 2).
FT	CONFLICT	132	132	N -> K (IN REF. 2).
FT	CONFLICT	172	172	S -> Y (IN REF. 2).
FT	CONFLICT	172	172	I -> V (IN REF. 2).
FT	CONFLICT	202	202	C -> V (IN REF. 2).
FT	SEQUENCE	269 AA;	28699 MW;	BACFG69AF4DDE56 CRC64;

CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Ala-|-Xaa. Does not
CC hydrolyse elastin.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC -!- CAUTION: Was originally (Ref.5) thought to be elastase 1.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M16630; AAA36482.1; .
CC EMBL: BC005216; AAH05216.1; .
CC EMBL: M18692; AA58454.1; .
CC PIR: B29934; B29934.
CC PIR: A27206; A27206.
CC PIR: S04999; S04999.
CC PIR: S04490; S04490.
CC HSP: P05805; IFON.
CC MEROPS: S01.205; .
CC GlycoSuiteDB; P08861; .
CC SWISS-2DPAGE; P08861; HUMAN.
CC Genew; HGNC:15945; ELA3B.
CC InterPro; IPR001314; Chymotrypsin.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; Chymotrypsin.
CC SMART; SM00020; TRYP_SPC; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC Hydrolase; Serine protease; Zymogen; Signal; Glycoprotein.
CC OR 16 (POTENTIAL).
CC SIGNAL 1 15
CC PROPEP 16 28
CC CHAIN 29 270
CC ELASTASE IIIB.
CC CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT_SITE 123 123
CC ACT_SITE 123 123
CC ACT_SITE 217 217
CC ACT_SITE 217 217
CC DISULFID 58 74
CC DISULFID 117 120
CC DISULFID 157 223
CC DISULFID 188 204
CC DISULFID 213 244
CC CARBOHYD 114 114
CC CONFLICT 4 4
CC CONFLICT 64 64
CC CONFLICT 79 79
CC CONFLICT 129 131
CC CONFLICT 164 164
CC CONFLICT 164 164
CC SEQUENCE 270 AA; 29293 MW; B14BE0AD3695AFE CRC64;
R -> G (IN REF. 3).
A -> G (IN REF. 3).
W -> R (IN REF. 1).
MISSING (IN REF. 5).
R -> P (IN REF. 3).
/FTIG-CAR_000212.
R -> G (IN REF. 3).
A -> G (IN REF. 3).
W -> R (IN REF. 1).
MISSING (IN REF. 5).
R -> P (IN REF. 3).
Query Match 37.1%; Score 453.5; DB 1; Length 270;
Best Local Similarity 42.5%; Pred. No. 3.5e-32;
Matches 102; Conservative 36; Mismatches 87; Indels 15; Gaps 8;
QY 2 LLSLTLVLSSGSGCGIPALPKALPSQRIVNGENAVLGSWPQVLSQ-DSSG--FH 58
DB 1 MMLRLSLLLAVASGYGPPSSRP-----SSRVNGEDAVPYSPWPQVLSQYKSGSFYH 56
QY 59 FCGSLISQSWVYTAACHNVPGRHFVILGEYDRSSNAEPLOVLSVSR--ITHPSWNST 116
DB 57 TCGSLIADPWVYTAGHICSSWTQVILGEYDRAVKEGPEQVPIPSGDLFVHPLWNKS 116
QY 117 TM--NNDVTLLKLSPAQYTRISPVLCASSNEALTEGLTCVTGTGWRSLSGVGNVTPAHL 174
DB 117 CVACGNDIALIKLSRAQLGDVAQLASLPAGDILNPTCYITGWRLYTNGPL-PDKL 175
QY 175 QVALPLVTNQCQR--YDSSITDSMICAGGAGSCCGDGGPLVC-QKGNWTWVLIGI 231
DB 176 QEALLPVVDYEHCSRNNWSSVYKTMVACAGGDIRSGCNGDGGGGLNCPTEDDGQVHGV 235

-RESULT 11
EL2_BOVIN STANDARD; PRT; 269 AA.
ID EL2_BOVIN
AC Q29461;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Elastase 2 precursor (EC 3.4.21.71).
GN ELA2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=98079203; PubMed=9418008;
RA Gestin M., le Huerou-Luron I., Wicker-Planquart C., le Drean G.,
RA Chaix J.C., Puigserver A., Guilloteau P.;
RT "Bovine pancreatic preproelastases I and II: comparison of nucleotide
RT and amino acid sequences and tissue specific expression.";
RL Comp. Biochem. Physiol. 118B:181-187 (1997).
CC -!- FUNCTION: ACTS UPON ELASTIN.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Leu-|-Xaa, Met-|-Xaa
CC and Phe-|-Xaa. Hydrolyzes elastin.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: PANCREAS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X97635; CAA66231.1; .
CC HSP: P00766; ICHG.
CC MEROPS: S01.155; .
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; Chymotrypsin.
CC SMART; SM00020; TRYP_SPC; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC Hydrolase; Serine protease; Zymogen; Signal.
CC SIGNAL 1 16
CC PROPEP 17 28
CC CHAIN 29 269
CC ELASTASE 2.
CC ACT_SITE 73 73
CC CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT_SITE 121 121
CC CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT_SITE 216 216
CC CHARGE RELAY SYSTEM (BY SIMILARITY).
CC DISULFID 58 74
CC DISULFID 155 222
CC DISULFID 186 202
CC DISULFID 212 243
CC SEQUENCE 269 AA; 28856 MW; 8343B97062CF267C CRC64;
Query Match 37.1%; Score 453; DB 1; Length 269;
Best Local Similarity 43.9%; Pred. No. 3.9e-32;
Matches 105; Conservative 30; Mismatches 90; Indels 14; Gaps 8;
QY 2 LLSLTLVLSSGSGCGIPALPKALPSQRIVNGENAVLGSWPQVLSQDSSG---FH 58
DB 1 MTRALLSLTVAG-ALSCGVPPYPPQLS---RVVGEDARPNWPQVLSQYSSSGQWRH 56
QY 59 FCGSLISQSWVYTAACHNVPGRHFVILGEYDRSSNAEPLOVLSVSR--ITHPSWNST 118


```
Db 57 TCGSLIEQNVLTAACHICSSRTYRVVGRQSLSTVSGSLTIAVSKSVIHEKWNSQL 116
Qy 119 --NDVTLKLLASPAQYTRISPVCLASNEALTEGLTCVTTWGRLSGVGNVTPAHLOQ 176
Db 117 AOGNDIALKLLASSVPLFDKIQGLCLPAAGTILPNNYCYVYTWGRLQSGNG-ALPDILQ 175
Qy 177 VALPLVTYNQCR--QYDSSITDSMICAGGAG-ASSCGDGGPLVCKGN-TWVLIGI 231
Db 176 GKLLVVDYATCSNFWGSTVKTNNICAGGDGVTSNCNGDGGPLNLCQANRQWQVHGI 234

RESULT 12
EL2A_HUMAN STANDARD; PRT; 269 AA.
AC P08217; Q14243;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Elastase 2A precursor (EC 3.4.21.71).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87217962; PubMed=3646943;
RA Kawashima I., Tani T., Shimoda K., Takiguchi Y.;
RT "Characterization of pancreatic elastase II cDNAs: two elastase II
RT mRNAs are expressed in human pancreas.";
RL DNA 6:163-172(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88107669; PubMed=3427074;
RA Fletcher T.S., Shen W.F., Largman C.;
RT "Primary structure of human pancreatic elastase 2 determined by
RT sequence analysis of the cloned mRNA.";
RL Biochemistry 26:7256-7261(1987).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=88198076; PubMed=2834346;
RA Shirasu Y., Yoshida H., Matsuki S., Takemura K., Ikeda N., Shimada Y.,
RA Ozawa T., Mikayama T., Iijima H., Ishida A., Sato Y., Tamai Y.,
RA Tanaka J., Ikenaga H.;
RT "Molecular cloning and expression in Escherichia coli of a cDNA
RT encoding human pancreatic elastase 2.";
RL J. Biochem. 102:1555-1563(1987).
RN [4]
RP SEQUENCE FROM N.A.
RA Thomas D.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ACTS UPON ELASTIN.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Leu-I-Xaa, Met-I-Xaa
CC and Phe-I-Xaa. Hydrolyzes elastin.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: PANCREAS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
Db EMBL; M16631; AAA52374.1; -;
```

```
DR EMBL; M16652; AAA52380.1; -;
DR EMBL; D00236; BAA00165.1; -;
DR EMBL; AL512883; CAC42421.1; -;
DR EMBL; BC007031; AA07031.1; -;
DR PIR; A27432; A27432.
DR PIR; B26823; B26823.
DR HSSP; P00772; 1ELG.
DR MEROPS; S01.155; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; zymogen; Signal.
FT SIGNAL 1 16
FT PROPEP 17 28
FT CHAIN 17 28 ACTIVATION PEPTIDE.
FT DISULFID 29 269
FT DISULFID 58 74
FT DISULFID 155 222 BY SIMILARITY.
FT DISULFID 186 202 BY SIMILARITY.
FT DISULFID 212 243 BY SIMILARITY.
FT ACT_SITE 73 73 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 121 121 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 216 216 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CONFLICT 202 202 C -> V (IN REF. 3).
SQ SEQUENCE 269 AA; 28888 MW; A2E05143EFF4987C CRC64;

Query Match 36.4%; Score 445; DB 1; Length 269;
Best Local Similarity 43.9%; Pred. No. 1.9e-31;
Matches 105; Conservative 28; Mismatches 92; Indels 14; Gaps 8;

Qy 2 LLSLTLSLVLLGSGWCGIPAIKPAISQRIVNGENAVLGSWPMQVSLQDSSG---FH 58
Db 1 MIRTLLLTSLVAG-ALSCGDPTYP---VYTRVVGEEARPNSWPQVSLQYSSNGKQW 56
Qy 59 FCGSLISQSWVYTAACHNVSPPGRHFVVLGEYDRSSNAEPLQVLSVRAITHPSNSTM 118
Db 57 TCGSLIANSVLTAACHICSSRTYRVGLGRHNLVVAESGLAVSVSVIVHDKWNSQI 116
Qy 119 N--NDVTLKLLASPAQYTRISPVCLASNEALTEGLTCVTTWGRLSGVGNVTPAHLOQ 176
Db 117 SKGNDIALKLANPVSLTDKIQGLCLPPAGTILPNNYCYVYTWGRLQNGAV-PDVLOQ 175
Qy 177 VALPLVTYNQCRQ---YDSSITDSMICAGGAG-SSCGDGGPLVCKGN-TWVLIGI 231
Db 176 GKLLVVDYATCSSAWWGSSVKTSMICAGGDGVTSNCNGDGGPLNLCQASDGRWQVHGI 234

RESULT 13
EL2_RAT ID EL2_RAT STANDARD; PRT; 271 AA.
AC P00774;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Elastase 2 precursor (EC 3.4.21.71).
GN ELA2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82182967; PubMed=6918221;
RA McDonald R.J., Swift G.H., Quinto C., Swain W., Pictet R.L.,
RA Nikovits W., Rutter W.J.;
RT "Primary structure of two distinct rat pancreatic preproelastases
RT determined by sequence analysis of the complete cloned messenger
RT ribonucleic acid sequences.";
RL Biochemistry 21:1453-1463(1982).
```

[illegible]

[illegible]

RESULT 15			
EL3A_HUMAN		EL3A_HUMAN	
ID	EL3A_HUMAN	STANDARD;	PRT; 270 AA.
AC	P09093; Q9BRW4;		
DT	01-MAR-1989 (Rel. 10, Created)		
DT	01-FEB-1996 (Rel. 33, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Elastase IIIA precursor (EC 3.4.21.70) (protease E).		

OC Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI TaxID=9606.

SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA MEDLINE=88087253; PubMed=2826474;
RX Tani T., Ohsumi J., Mita K., Takiguchi Y.;
RT "Identification of a novel class of elastase isozyme, human
RL pancreatic elastase III, by cDNA and genomic gene cloning.";
RL J. Biol. Chem. 263:1231-1239(1988).

RC TISSUE-Pancreas;[14] AND ENKALIN SEQUENCE.
RX MEDLINE-89303417; Pubmed-2460440;
RA Shirasu Y., Takemura K., Yoshida H., Sato Y., Iijima H.,
RA Shimada Y., Mikayama T., Ozawa T., Ikeda N., Ishida A., Tamai Y.,
RA Matsuki S., Tanaka J., Ikenaga H., Ogawa M.;
RT "Molecular cloning of complementary DNA encoding one of the human
RT pancreatic protease E isozymes.";
RL J. Biochem. 104:259-264(1988).

RC TISSUE=prostate;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: EFFICIENT PROTEASE WITH ALANINE SPECIFICITY BUT ONLY
CC LITTLE ELASTOLYTIC ACTIVITY.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Ala-I-Xaa. Does not
CC hydrolyse elastin.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

DR	EMBL; M18700; AAA66350.1; .	JOINED.
DR	EMBL; M18693; AAA66350.1; .	JOINED.
DR	EMBL; M18694; AAA66350.1; .	JOINED.
DR	EMBL; M18695; AAA66350.1; .	JOINED.
DR	EMBL; M18696; AAA66350.1; .	JOINED.
DR	EMBL; M18697; AAA66350.1; .	JOINED.
DR	EMBL; M18698; AAA66350.1; .	JOINED.
DR	EMBL; M18699; AAA66350.1; .	JOINED.
DR	EMBL; M18699; AAA66350.1; .	JOINED.
DR	EMBL; D00306; BAA00212.1; -.	
DR	EMBL; BC005918; AAH05918.1; -.	
DR	PIR; A29934; A29934.	
DR	HSSP; P05805; 1FON.	
DR	MEROPS; S01.154; -.	

DR	Genew; HGNC:15944; ELA3A.	
DR	InterPro: IPR001314; Chymotrypsin.	
DR	InterPro: IPR001254; Ser_protease_Try.	
DR	Pfam: PF00089; trypsin; 1.	
DR	PRINTS: PR00722; CHYMOTRYPSIN.	
DR	SMART: SM00020; TRYD_SPC; 1.	
DR	PROSITE: PS50240; TRYPSIN_DOM; 1.	
DR	PROSITE: PS00134; TRYPSIN_HIS; 1.	
DR	PROSITE: PS00135; TRYPSIN_SER; 1.	
DR	KW Hydrolase; Serine protease; Zymogen; Signal; Glycoprotein.	
FT	SIGNAL 1 15	(POTENTIAL).
FT	PROPEP 16 28	ACTIVATION PEPTIDE (POTENTIAL).
FT	CHAIN 29 270	ELASTASE IIIA.
FT	ACT_SITE 73 73	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 123 123	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 217 217	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	DISULFID 58 74	BY SIMILARITY.
FT	DISULFID 117 120	PROBABLE.
FT	DISULFID 157 223	BY SIMILARITY.
FT	DISULFID 188 204	BY SIMILARITY.
FT	DISULFID 213 244	BY SIMILARITY.
FT	CARBOHYD 114 114	N-LINKED (GLCNAC...) (POTENTIAL).
FT	CONFLICT 63 63	I -> T (IN REF. 3).
FT	CONFLICT 106 106	MISSING (IN REF. 1).
FT	CONFLICT 174 174	K -> E (IN REF. 3).
FT	SEQUENCE 270 AA; 29474 MW; 576DDB255A4118C CRC64;	
SO		

Search completed: December 20, 2002, 15:17:12
Job time : 8.91304 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:16:49 ; Search time 1.04348 seconds
(without alignments)
530.147 Million cell updates/sec

Title: US-09-856-319B-4_COPY_1_33

Perfect score: 166

Sequence: 1 MLLSLTSLVLLGSSWCGVPAITPALSYNQR 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 106657 seqs, 16763532 residues

Total number of hits satisfying chosen parameters: 106657

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	133	80.1	192	10	US-09-925-297-529
2	68	41.0	263	10	US-09-888-615-96
3	67	40.4	146	10	US-09-925-297-643
4	55	33.1	231	9	US-10-095-449-2
5	55	33.1	231	10	US-09-448-378-2
6	55	33.1	231	10	US-09-983-806-2
7	55	33.1	242	10	US-09-815-242-13730
8	50	30.1	344	10	US-09-992-647-12
9	48	28.9	13	12	US-10-036-371-5
10	48	28.9	331	9	US-09-860-670-118
11	48	28.9	331	10	US-09-764-853-671
12	48	28.9	331	10	US-09-764-898-219
13	48	28.9	652	10	US-09-992-647-1
14	48	28.9	661	10	US-09-764-853-679
15	48	28.9	661	10	US-09-764-898-224
16	47.5	28.6	20	9	US-10-010-928-7
17	47.5	28.6	20	10	US-09-802-124-7
18	47.5	28.6	20	10	US-09-733-605-7
19	47.5	28.6	78	9	US-09-992-598-363

20	47.5	28.6	78	9	US-09-989-293A-363	Sequence 363, App
21	47.5	28.6	78	9	US-10-063-547-66	Sequence 66, Appl
22	47.5	28.6	78	9	US-09-092-296-15	Sequence 15, Appl
23	47.5	28.6	78	10	US-09-989-722-363	Sequence 363, App
24	47.5	28.6	78	10	US-09-989-723-363	Sequence 363, App
25	47.5	28.6	78	10	US-09-989-279-363	Sequence 363, App
26	47.5	28.6	78	10	US-09-989-721-363	Sequence 363, App
27	47.5	28.6	78	10	US-09-989-731-363	Sequence 363, App
28	47.5	28.6	78	10	US-09-989-732-363	Sequence 363, App
29	47.5	28.6	78	10	US-09-991-073-363	Sequence 363, App
30	47.5	28.6	78	10	US-09-990-442-363	Sequence 363, App
31	47.5	28.6	78	10	US-09-991-163-363	Sequence 363, App
32	47.5	28.6	78	10	US-09-993-604-363	Sequence 363, App
33	47.5	28.6	78	10	US-09-990-456-363	Sequence 363, App
34	47.5	28.6	78	10	US-09-989-721-363	Sequence 363, App
35	47.5	28.6	78	12	US-10-006-867-66	Sequence 66, Appl
36	47.5	28.6	78	12	US-10-052-586-242	Sequence 242, App
37	47.5	28.6	180	9	US-10-010-928-2	Sequence 2, Appl
38	47.5	28.6	180	10	US-09-802-124-2	Sequence 2, Appl
39	47.5	28.6	180	10	US-09-733-605-2	Sequence 2, Appl
40	47	28.3	237	10	US-09-746-284-1	Sequence 1, Appl
41	47	28.3	413	10	US-09-815-242-13998	Sequence 13998, A
42	47	28.3	445	9	US-09-992-598-177	Sequence 177, App
43	47	28.3	445	9	US-09-989-293A-177	Sequence 177, App
44	47	28.3	445	9	US-10-063-547-32	Sequence 32, Appl
45	47	28.3	445	10	US-09-989-722-177	Sequence 177, App

ALIGNMENTS

RESULT 1
US-09-925-297-529
; Sequence 529, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 529
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-297-529

Query Match 80.1%; Score 133; DB 10; Length 192;
Best Local Similarity 86.2%; Pred. No. 2e-10;
Matches 25; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 SUTLSVLVLGSSWCGVPAITPALSYNQR 33
|||||
Db 1 SUTLSVLVLGSSWCGVPAITPALSFSQR 29

RESULT 2
US-09-888-615-96
; Sequence 96, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA

```

; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038600/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 96
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-96

```

```
Query Match      41.0%; Score 68; DB 10; Length 263;
Best Local Similarity 51.7%; Pred. No. 0.073;
Matches 15; Conservative 4; Mismatches 10; Indels 0; Gaps 0;
```

QY 1 MLLSLTSLVLLGSSWGCGVPATPALS 29
| | | | | : : : : :
Db 1 MASLWLLSCFSLVGAAGCGVPALHPVLS 29

```

RESULT 3
US - 09-925-297-643
, Sequence 643, Application US/09925297
, Patent No. US20020081659A1
, GENERAL INFORMATION:
, APPLICANT: Rosen et al.
, TITLE OF INVENTION: Nucleic Acids, P
, FILE REFERENCE: PA105
, CURRENT APPLICATION NUMBER: US/09/925
, CURRENT FILING DATE: 2001-08-10
, PRIOR APPLICATION NUMBER: PCT/US00/08
, PRIOR FILING DATE: 2000-03-08
, PRIOR APPLICATION NUMBER: 60/124,270
, PRIOR FILING DATE: 1999-03-12
, NUMBER OF SEQ ID NOS: 928
, SOFTWARE: PatentIn Ver. 2.0
, SEQ ID NO 643
, LENGTH: 146

```

```

/ ORGANISM: Homo sapiens
/
/ FEATURE:
/
/ NAME/KEY: SITE
/ LOCATION: (2)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/
/ NAME/KEY: SITE
/ LOCATION: (94)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/
/ NAME/KEY: SITE
/ LOCATION: (126)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/
/ NAME/KEY: SITE
/ LOCATION: (130)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/
/ NAME/KEY: SITE
/ LOCATION: (133)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/
/ NAME/KEY: SITE
/ LOCATION: (137)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/
/ NAME/KEY: SITE
/ LOCATION: (143)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/
/ US-925-297-643

```

Query Match 40.4%; Score 67; DB 10; Length 146;
Best Local Similarity 66.7%; Pred. No. 0.053;
Matches 12; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 12 LLGSSWCGVPAITPALS 29
| : : : | | | | |

Db 16 LVGAFCGVPAIHPVLS 33

RESULT 4

US-10-095-449-2
; Sequence 2, Application US/10095449
; Patent No. US20020160004A1
; GENERAL INFORMATION:

GENERAL INFORMATION:
APPLICANT: Lyman, Stewart D.
Beckmann, M. Patricia
TITLE OF INVENTION: Ligands for flt3/flk-2 Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stephen L. Malaska, Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: US

COMPUTED READABLE FORM:
T0196 :PT7

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0.1
SOFTWARE: Microsoft Word, Version #5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/095.449
FILING DATE: 13-Mar-2002
CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/669,692
FILING DATE: 24-JUN-1996
APPLICATION NUMBER: US/08/162,407
FILING DATE: December 3, 1993
APPLICATION NUMBER: 08/111,758
FILING DATE: August 25, 1993
APPLICATION NUMBER: 08/106,463
FILING DATE: August 12, 1993
APPLICATION NUMBER: 08/068,394
FILING DATE: May 24, 1993

ATTORNEY/AGENT INFORMATION:
NAME: Malaska, Stephen L.
REGISTRATION NUMBER: 32,655
REFERENCE/DOCKET NUMBER: 2813-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822

INFORMATION FOR SEO ID NO: 2:
1 FEB 19 1962

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 231 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-095-449-2

```

Query Match 33.1%; Score 55; DB 9; Length 231;
Best Local Similarity 61.1%; Pred. No. 3.1;
Matches 11; Conservative 4; Mismatches 3; Indels

Qy 1 MLLSLTSLVLLGSSWG 18
:||| | | :||| :|||
Db 192 LLLLLPLTLVLLAAWG 209

RESULTS

```

US-09-448-378-2
; Sequence 2, Application US/09448378
; Patent No. US20020034517A1
; GENERAL INFORMATION:
; APPLICANT: Brasel, Kenneth
; TITLE OF INVENTION: Dendritic Ce
; FILE REFERENCE: 2836-D

```

Query Match 33.1%; Score 55; DB 10; Length 231;
Best Local Similarity 61.1%; Pred. No. 3.1;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
CURRENT APPLICATION NUMBER: US/09/448,378
CURRENT FILING DATE: 1999-11-23
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 231
TYPE: PRT
ORGANISM: Mus sp.
US-09-448-378-2

Qy 1 MLLSLTSLVLGSSWG 18
:|||||:|||||:|||||
Db 192 LLLLLPLTLVLLAAWG 209

RESULT 6

US-09-983-806-2
Sequence 2, Application US/09983806
Patent No. US20020107365A1
GENERAL INFORMATION:
APPLICANT: Lyman, Stewart D.
Beckmann, M. Patricia
TITLE OF INVENTION: Ligands for flk3/flk-2 Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stephen L. Malaska, Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0.1
SOFTWARE: Microsoft Word, Version #5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/983,806
FILING DATE: 25-Oct-2001
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/444,626
FILING DATE: 19-MAY-1995
APPLICATION NUMBER: US/08/162,407
FILING DATE: 03-DEC-1993
APPLICATION NUMBER: 08/111,758
FILING DATE: August 25, 1993
APPLICATION NUMBER: 08/106,463
FILING DATE: August 12, 1993
APPLICATION NUMBER: 08/068,394
FILING DATE: May 24, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Malaska, Stephen L.
REGISTRATION NUMBER: 32,655
REFERENCE/DOCKET NUMBER: 2813-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 231 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-983-806-2

Query Match 33.1%; Score 55; DB 10; Length 231;

Best Local Similarity 61.1%; Pred. No. 3.1;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
Qy 1 MLLSLTSLVLGSSWG 18
:|||||:|||||:|||||
Db 192 LLLLLPLTLVLLAAWG 209

RESULT 7

US-09-815-242-13730
Sequence 13730, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13730
LENGTH: 242
TYPE: PRT
ORGANISM: Salmonella typhi
US-09-815-242-13730

Query Match 33.1%; Score 55; DB 10; Length 242;
Best Local Similarity 58.8%; Pred. No. 3.3;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 17 WCGGVPATPALSYNOR 33
||| ||||| |||||
Db 92 WGFDPVPAIEKALAHNER 108

RESULT 8

US-09-992-647-12
Sequence 12, Application US/09992647
Patent No. US20020146767A1
GENERAL INFORMATION:
APPLICANT: Xu, Hong
APPLICANT: Cohan, Victoria L.
APPLICANT: Stuart, Susan G.
TITLE OF INVENTION: HUMAN EMR1-LIKE G PROTEIN-COUPLED RECEPTOR
FILE REFERENCE: PC-0052 CIP
CURRENT APPLICATION NUMBER: US/09/992,647
CURRENT FILING DATE: 2002-11-13
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PERL Program
SEQ ID NO 12
LENGTH: 344

```

RESULT 11
US-09-764-853-671
; Sequence 671, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P0206
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 671
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-853-671

Query Match      28.9%; Score 48; DB 10; Length 331;
Best Local Similarity 37.0%; Pred. No. 37;
Matches 10; Conservative 4; Mismatches 13; Indels

QY      1 MLLSLTSLVLLGSSWCGVPAITPA 27
      || | | | | | | | | | | | | | |
DB      228 MIAFKATAQLFILGCTWCLGLLQVGPA 254

RESULT 12
US-09-764-898-219
; Sequence 219, Application US/09764898
; Patent No. US20020090673A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P0201
; CURRENT APPLICATION NUMBER: US/09/764,898
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 311
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 219
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-898-219

Query Match      28.9%; Score 48; DB 10; Length 331;
Best Local Similarity 37.0%; Pred. No. 37;
Matches 10; Conservative 4; Mismatches 13; Indels

QY      1 MLLSLTSLVLLGSSWCGVPAITPA 27
      || | | | | | | | | | | | | | |
DB      228 MIAFKATAQLFILGCTWCLGLLQVGPA 254

RESULT 13
US-09-992-647-1
; Sequence 1, Application US/09992647
; Patent No. US20020146767A1
; GENERAL INFORMATION:
; APPLICANT: Cohan, Victoria L.
; APPLICANT: Stuart, Susan G.
; TITLE OF INVENTION: HUMAN EMR1-LIKE G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCE: PC-0052 CIP
; CURRENT APPLICATION NUMBER: US/09/992,647
; CURRENT FILING DATE: 2002-11-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PERL Program
; SEQ ID NO 1

```


Search completed: December 20, 2002, 15:21:07
Job time : 1.04348 secs

```
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID NO. US20020146767A1 429905
US-09-992-647-1

Query Match      28.9%; Score 48; DB 10; Length 652;
Best Local Similarity 37.0%; Pred. No. 75;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLLGSSWCGVPAITPA 27
   || | | | | | | | | | |
DB 549 MIAFRATAQLFILGCTWCLGLLQVGPA 575

RESULT 14
US-09-764-853-679
; Sequence 679, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJZ06
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 679
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-853-679

Query Match      28.9%; Score 48; DB 10; Length 661;
Best Local Similarity 37.0%; Pred. No. 76;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLLGSSWCGVPAITPA 27
   || | | | | | | | | | |
DB 558 MIAFRATAQLFILGCTWCLGLLQVGPA 584

RESULT 15
US-09-764-898-224
; Sequence 224, Application US/09764898
; Patent No. US20020090673A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJZ01
; CURRENT APPLICATION NUMBER: US/09/764,898
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 311
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 224
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-898-224

Query Match      28.9%; Score 48; DB 10; Length 661;
Best Local Similarity 37.0%; Pred. No. 76;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLLGSSWCGVPAITPA 27
   || | | | | | | | | | |
DB 558 MIAFRATAQLFILGCTWCLGLLQVGPA 584
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:14:53 : Search time 1.86957 Seconds
(without alignments)
519.349 Million cell updates/sec

Title: US-09-856-319B-4_COPY_1_33

Perfect score: 166

Sequence: 1 MLLSLTSLVLLGSSWCGVPAITPALSYNQR 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*

2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*

3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*

4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*

5: /cgn2_6/ptodata/1/1aa/PCRUS.COMB.pep.*

6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55.5	33.4	441	3	US-09-052-778-14
2	55	33.1	231	1	US-08-220-379B-7
3	55	33.1	231	1	US-08-243-545-2
4	55	33.1	231	2	US-08-993-962-2
5	55	33.1	231	4	US-09-160-841-2
6	55	33.1	231	5	PCT-US94-05365-2
7	55	33.1	231	5	PCT-US95-03866-6
8	55	33.1	268	1	US-08-270-584A-2
9	55	33.1	268	3	US-08-765-192-2
10	55	33.1	268	3	US-09-199-793-2
11	52	31.3	306	4	US-09-386-642-53
12	52	31.3	319	4	US-09-386-642-12
13	50	30.1	344	3	US-09-110-116-4
14	48.5	29.2	280	4	US-09-296-284-27
15	48.5	29.2	306	4	US-09-296-284-6
16	48	28.9	521	4	US-08-956-322-4
17	48	28.9	652	3	US-09-110-116-1
18	48	28.9	652	4	US-08-956-322-2
19	47	28.3	838	2	US-08-696-944-19
20	47	28.3	1620	1	US-08-542-363-2
21	47	28.3	1620	4	US-09-100-089-2
22	47	28.3	1620	4	US-09-670-827-2
23	46.5	28.0	428	3	US-09-052-778-12
24	46.5	28.0	586	4	US-09-134-001C-4456
25	46	27.7	147	4	US-09-134-001C-3180
26	46	27.7	365	2	US-08-855-518-5
27	46	27.7	366	2	US-08-855-518-4

Sequence 2, Appli
Sequence 14, Appl
Sequence 16, Appl
Sequence 18, Appl
Sequence 20, Appl
Sequence 22, Appl
Sequence 24, Appl
Sequence 26, Appl
Sequence 28, Appl
Sequence 30, Appl
Sequence 32, Appl
Sequence 34, Appl
Sequence 36, Appl
Sequence 38, Appl
Sequence 40, Appl
Sequence 42, Appl
Sequence 44, Appl
Sequence 46, Appl

28 46 27.7 553 2 US-08-943-087-2
29 46 27.7 553 2 US-08-943-087-14
30 46 27.7 553 2 US-08-943-087-16
31 46 27.7 553 2 US-08-943-087-18
32 46 27.7 553 2 US-08-943-087-20
33 46 27.7 553 2 US-08-943-087-22
34 46 27.7 553 2 US-08-943-087-24
35 46 27.7 553 2 US-08-943-087-26
36 46 27.7 553 2 US-08-943-087-28
37 46 27.7 553 2 US-08-943-087-30
38 46 27.7 553 2 US-08-943-087-32
39 46 27.7 553 2 US-08-943-087-34
40 46 27.7 553 2 US-08-943-087-36
41 46 27.7 553 2 US-08-943-087-38
42 46 27.7 553 2 US-08-943-087-40
43 46 27.7 553 2 US-08-943-087-42
44 46 27.7 553 2 US-08-943-087-44
45 46 27.7 553 2 US-08-943-087-46

ALIGNMENTS

RESULT 1

US-09-052-778-14
: Sequence 14, Application US/09052778A
: Patent No. 6060590
: GENERAL INFORMATION:
: APPLICANT: Bryant, Peter J.
: APPLICANT: Kawamura, Kazuo
: TITLE OF INVENTION: CHITINASE RELATED PROTEINS AND METHODS
: TITLE OF INVENTION: OF USE
: FILE REFERENCE: 07306/015001
: CURRENT APPLICATION NUMBER: US/09/052,778A
: CURRENT FILING DATE: 1998-03-31
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 14
: LENGTH: 441
: TYPE: PRT
: ORGANISM: Drosophila melanogaster
US-09-052-778-14

Query Match 33.4%; Score 55.5; DB 3; Length 441;
Best Local Similarity 48.3%; Pred. No. 8.3;
Matches 14; Conservative 5; Mismatches 9; Indels 1; Gaps 1;
Qy 2 LLSLTSLVLLGSSWCGVPAITPALSY 30
|||||: | : ||| | : | : | :
Db 210 LLSLTIV-LPNVNSSWYDAPSIAPSLDF 237

RESULT 2

US-08-220-379B-7
: Sequence 7, Application US/08220379B
: Patent No. 5525708
: GENERAL INFORMATION:
: APPLICANT: No. 5525708ka, Karl
: APPLICANT: Lobell, Robert B
: TITLE OF INVENTION: STABILIZED DIMER OF KIT LIGAND
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Neave
: STREET: 1251 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: United States of America
: ZIP: 10020
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/220,379B
; FILING DATE: 28-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: CytoMed/2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 231 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; PUBLICATION INFORMATION:
; AUTHORS: Lyman, et al.
; JOURNAL: Cell
; PAGES: 1157-1167
; DATE: 1993
; RELEVANT RESIDUES IN SEQ ID NO: 7: FROM 1 TO 231
US-08-220-379B-7

Query Match 33.1%; Score 55; DB 1; Length 231;
Best Local Similarity 61.1%; Pred. No. 4.7;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGSSWG 18
:|||||:||||:|
Db 192 LLLLLPLTLVLAAGW 209

RESULT 3
US-08-243-545-2
; Sequence 2, Application US/08243545
; Patent No. 5534512
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; APPLICANT: Beckmann, M. Patricia
; TITLE OF INVENTION: Ligands for flt3/flk-2 Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Stephen L. Malaska, Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.0.1
; SOFTWARE: Microsoft Word, Version #5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/243,545
; FILING DATE: 11-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/162,407
; FILING DATE: 03-DEC-1993
; APPLICATION NUMBER: 08/111,758
; FILING DATE: August 25, 1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/106,463
; FILING DATE: August 12, 1993
; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/068,394
; FILING DATE: May 24, 1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Malaska, Stephen L.
; REGISTRATION NUMBER: 32,655
; REFERENCE/DOCKET NUMBER: 2813-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; TELEX: 756822
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 231 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-243-545-2

Query Match 33.1%; Score 55; DB 1; Length 231;
Best Local Similarity 61.1%; Pred. No. 4.7;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGSSWG 18
:|||||:||||:|
Db 192 LLLLLPLTLVLAAGW 209

RESULT 4
US-08-993-962-2
; Sequence 2, Application US/08993962
; Patent No. 5843423
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; APPLICANT: Beckmann, M. Patricia
; TITLE OF INVENTION: Ligands for flt3/flk-2 Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Stephen L. Malaska, Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.0.1
; SOFTWARE: Microsoft Word, Version #5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,962
; FILING DATE: December 18, 1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/162,407
; FILING DATE: December 3, 1993
; APPLICATION NUMBER: 08/111,758
; FILING DATE: August 25, 1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/106,463
; FILING DATE: August 12, 1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/068,394
; FILING DATE: May 24, 1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Malaska, Stephen L.
; REGISTRATION NUMBER: 32,655
; REFERENCE/DOCKET NUMBER: 2813-C
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 231 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-993-962-2

Query Match 33.1%; Score 55; DB 2; Length 231;
Best Local Similarity 61.1%; Pred. No. 4.7;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MLLSLTSLVLLGSSWG 18
:|||||:|||||:
Db 192 LLLLLPLTLVLLAAWG 209

RESULT 5

US-09-160-841-2
Sequence 2, Application US/09160841
Patent No. 6190655
GENERAL INFORMATION:
APPLICANT: Lyman, Stewart D.
APPLICANT: Beckmann, M. Patricia
TITLE OF INVENTION: Ligands for flt3/flk-2 Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stephen L. Malaska, Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0.1
SOFTWARE: Microsoft Word, Version #5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/160,841
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/162,407
FILING DATE: December 3, 1993
APPLICATION NUMBER: 08/111,758
FILING DATE: August 25, 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/106,463
FILING DATE: August 12, 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/068,394
FILING DATE: May 24, 1993

ATTORNEY/AGENT INFORMATION:
NAME: Malaska, Stephen L.
REGISTRATION NUMBER: 32,655
REFERENCE/DOCKET NUMBER: 2813-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 231 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-160-841-2

Query Match 33.1%; Score 55; DB 4; Length 231;

Best Local Similarity 61.1%; Pred. No. 4.7;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MLLSLTSLVLLGSSWG 18
:|||||:|||||:
Db 192 LLLLLPLTLVLLAAWG 209

RESULT 6

PCT-US94-05365-2
Sequence 2, Application PC/TUS9405365
GENERAL INFORMATION:
APPLICANT: Lyman, Stewart D.
APPLICANT: Beckmann, M. Patricia
TITLE OF INVENTION: Ligands for flt3/flk-2 Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stephen L. Malaska, Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05365
FILING DATE: May 24, 1994
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: -to be assigned-
FILING DATE: May 11, 1994
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/209,502
FILING DATE: March 7, 1994
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/162,407
FILING DATE: December 3, 1993
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/111,758
FILING DATE: August 25, 1993
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/106,463
FILING DATE: August 12, 1993
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/068,394
FILING DATE: May 24, 1993
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Malaska, Stephen L.
REGISTRATION NUMBER: 32,655
REFERENCE/DOCKET NUMBER: 2813-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 231 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-05365-2

Query Match 33.1%; Score 55; DB 5; Length 231;

Best Local Similarity 61.1%; Pred. No. 4.7;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLLGSSWG 18
:|||||:|||||:|
Db 192 LLLLLPLTLVLLAAWG 209

RESULT 7

PCT-US95-03866-6
; Sequence 6, Application PC/TUS9503866
; GENERAL INFORMATION:
; APPLICANT: CytoMed, Inc. (all states except US)
; APPLICANT: Nocka, Karl (US only)
; APPLICANT: Lobell, Robert B (US only)
; TITLE OF INVENTION: STABILIZED DIMER OF KIT LIGAND AND
; TITLE OF INVENTION: FLT-3/FLK-2 LIGAND
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/03866
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,379
; FILING DATE: 28-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr, James F
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: CytoMed/2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 231 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..205
PCT-US95-03866-6

Query Match 33.1%; Score 55; DB 5; Length 231;
Best Local Similarity 61.1%; Pred. No. 4.7;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLLGSSWG 18
:|||||:|||||:|
Db 192 LLLLLPLTLVLLAAWG 209

RESULT 8

US-08-270-584A-2
; Sequence 2, Application US/08270584A
; Patent No. 5710035
; GENERAL INFORMATION:
; APPLICANT: GREENE, ET AL.
; TITLE OF INVENTION: Human Elastase IV

; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068

; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/270,584A
; FILING DATE: July 5, 1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FERRARO, GREGORY D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-126
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 268 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; US-08-270-584A-2

Query Match 33.1%; Score 55; DB 1; Length 268;
Best Local Similarity 40.7%; Pred. No. 5.5;
Matches 11; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 3 LLSLTSLVLLGSSWCGVPATPALS 29
:|:|:|:|:|:|:|:|:|:|
Db 1 MLGITVLAALLACASSCGVPSPPNLS 27

RESULT 9

US-08-765-192-2
; Sequence 2, Application US/08765192
; Patent No. 5851814
; GENERAL INFORMATION:
; APPLICANT: Greene, John et al.
; TITLE OF INVENTION: Human Elastase IV
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/765,192
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373

Db 1 MLGITVLAALLACASSCGVPSEFPNLS 27

```
; ORGANISM: Gluconobacter suboxydans
US-09-296-284--6

Query Match          29.2%; Score 48.5; DB 4; Length 306;
Best Local Similarity 45.5%; Pred. No. 48;
Matches 10; Conservative 3; Mismatches 8; Indels 1; Gaps 1;

QY      6 LTLSLVL-LGSSWCGVPAITP 26
       :| | | | | | | | | |
Db     269 IRLALALARKWGAVPKVLP 290
```

Search completed: December 20, 2002, 15:20:38
Job time : 2.86957 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:12:13 ; Search time 5.08696 Seconds
(without alignments)
864.421 Million cell updates/sec

Title: US-09-856-319B-4_COPY_1_33

Perfect score: 166

Sequence: 1 MLLSLTSLVLLGSSWCGVPAITPALSYNQR 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	166	100.0	264	21 AAB11711	Mouse serine prote
2	150	90.4	264	21 AAB11710	Human serine prote
3	133	80.1	192	21 AAB54077	Human pancreatic c
4	68	41.0	44	21 AAY64742	Human 5' EST relat
5	68	41.0	263	23 AAU82738	Amino acid sequenc
6	67	40.4	146	21 AAB54191	Human pancreatic c
7	59	35.5	195	20 AAV11649	Human 5' EST secre
8	58	34.9	306	23 AAU95588	Human olfactory an
9	58	34.9	312	22 AAG71664	Human olfactory re
10	58	34.9	312	22 AAG72192	Human olfactory re

11	58	34.9	312	22 AAG72355	Human OR-like poly
12	55.5	33.4	441	21 AAB07181	Drosophila melanog
13	55.5	33.4	441	22 AAB60694	Drosophila melanog
14	55	33.1	123	20 AAY11650	Human 5' EST secre
15	55	33.1	231	16 AAB67540	Mouse flt-3 ligand
16	55	33.1	231	20 AAB67768	Murine flt3-ligand
17	55	33.1	231	22 AAB20186	Mouse Flt-3 ligand
18	55	33.1	232	16 AAB66177	Mouse Mott10/T118
19	55	33.1	232	22 AAB20189	Mouse Flt-3 ligand
20	55	33.1	242	22 AAU38137	Salmonella typhi c
21	55	33.1	268	17 AAR90683	Human caldesmon co
22	55	33.1	268	17 AAR88481	Human elastase IV
23	54	32.5	13	20 AAY50212	Neutrophil-activat
24	54	32.5	84	22 AAU31895	Novel human secret
25	52.5	31.6	760	22 AAB66394	Human prostate ECG
26	52	31.3	55	22 AAB67538	Amino acid sequenc
27	52	31.3	66	22 AAB67539	Amino acid sequenc
28	52	31.3	222	20 AAV37759	Amino acid sequenc
29	52	31.3	306	22 AAB67544	Amino acid sequenc
30	52	31.3	319	21 AAB36481	Fusion gene with h
31	52	31.3	319	22 AAB67541	Amino acid sequenc
32	51	30.7	220	21 AAY73464	Human secreted pro
33	51	30.7	306	20 AAY41332	Human polyepitide
34	51	30.7	478	23 AAB90142	Human protein sequ
35	51	30.7	484	22 AAB93904	Herbicidally activ
36	51	30.7	826	23 ABB91392	Human ORFX ORF1716
37	50.5	30.4	296	21 AAB41952	Drosophila melanog
38	50	30.1	144	22 ABB67458	Human prostate tum
39	50	30.1	162	20 AAY73942	Rat camello 1 (Rcm
40	50	30.1	222	22 AAB19992	Human cell surface
41	50	30.1	453	21 AAY94340	Novel human diagno
42	50	30.1	517	22 ABB14307	Human mature EGF-1
43	50	30.1	800	23 AAE13609	Human EGF-like mol
44	50	30.1	823	23 AAE13608	Nucleus specific
45	49.5	29.8	135	19 AAW56733	

ALIGNMENTS

RESULT 1
AAB11711
ID AAB11711 standard; Protein: 264 AA.
AC AAB11711;
XX
XX 23-OCT-2000 (first entry)
DT
XX
XX Mouse serine protease BSSP5 (mbSSP5) SEQ ID NO:4.

BSSP5: serine protease; human; hbSSP5: mouse; mbSSP5: brain;
diagnostic marker; antibody; transgenic animal; Alzheimer's disease;
epilepsy; cancer; inflammation; infertility; pancreatitis;
prostatic hypertrophy.

Mus sp.

WO200031243-A1.

02-JUN-2000.

19-NOV-1999; 99WO-JP06473.

20-NOV-1998; 98JP-0347806.

(FUSO) FUSO PHARM IND LTD.

Uemura H, Okui A, Kominami K, Yamaguchi N, Mitsui S;

WPI: 2000-400058/34.

N-PSDB: AAA61734.

Serine proteases BSSP5, useful in detecting homologs, mutants and

polymorphic variants as markers for diagnosis of e.g. Alzheimer's disease, epilepsy, cancer and inflammation, using blood, urine, pancreas or other tissues -

Claim 3: Page 55-56; 70pp; Japanese.

The invention relates to novel serine proteases designated BSSP5 (AAB11710-B11711), and to nucleic acids encoding them (AAA61733-A61734). The invention also relates to vectors and transformants comprising BSSP5 nucleic acids; transgenic animals in which the expression level of BSSP5 can be varied; and an mBSSP5 knockout mouse. The invention additionally encompasses anti-BSSP5 antibodies and methods of production of such antibodies, methods of BSSP5 detection using the antibodies, and the use of BSSP5 proteins or fragments as diagnostic markers for certain medical conditions, e.g., pancreatitis. A method for detecting pancreaticitis comprising measuring BSSP5 concentration in the blood or urine, and a pancreatitis diagnostic agent containing an anti-BSSP5 antibody is also disclosed. Nucleotides encoding BSSP5 were initially isolated in a human brain cDNA library using degenerate PCR primers (AAA61744-A61745) based on conserved regions of serine proteases. The BSSP5 serine proteases and nucleotides encoding them are useful in detecting homologues, mutants and polymorphic variants in biological samples (e.g., blood, urine, brain, prostate gland, placenta, testis, pancreas and spleen) as diagnostic markers for conditions such as Alzheimer's disease, epilepsy, cancer, inflammation, infertility, pancreatitis and prostatic hypertrophy. Sequence AAB11710 represents human BSSP5 (hBSSP5), and sequence AAB11711 represents murine BSSP5 (mBSSP5).

Sequence 264 AA;

Query Match 100.0%; Score 166; DB 21; Length 264;
Best Local Similarity 100.0%; Pred. No. 1e-13;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGGSGGCVPAITPALSYNQR 33
|||||
DB 1 MLLSLTSLVLGGSGGCVPAITPALSYNQR 33

RESULT 2

AAB11710
ID AAB11710 standard; Protein; 264 AA.

AC AAB11710;

DT 23-OCT-2000 (first entry)

DE Human serine protease BSSP5 (hBSSP5) SEQ ID NO:2.

KW BSSP5; serine protease; human; hBSSP5; mouse; mBSSP5; brain;
diagnostic marker; antibody; transgenic animal; Alzheimer's disease;
epilepsy; cancer; inflammation; infertility; pancreatitis;
prostatic hypertrophy.

OS Homo sapiens.

PN WO200031243-A1.

PD 02-JUN-2000.

PF 19-NOV-1999; 99WO-JP06473.

PR 20-NOV-1998; 98JP-0347806.

PA (FUSO) FUSO PHARM IND LTD.

PI Uemura H, Okui A, Kominami K, Yamaguchi N, Mitsui S;

DR WPI: 2000-400058/34.

DR N-PSDB; AAA61733.

XX Serine proteases BSSP5, useful in detecting homologs, mutants and

polymorphic variants as markers for diagnosis of e.g. Alzheimer's disease, epilepsy, cancer and inflammation, using blood, urine, pancreas or other tissues -

Claim 1: Page 51-52; 70pp; Japanese.

The invention relates to novel serine proteases designated BSSP5 (AAB11710-B11711), and to nucleic acids encoding them (AAA61733-A61734). The invention also relates to vectors and transformants comprising BSSP5 nucleic acids; transgenic animals in which the expression level of BSSP5 can be varied; and an mBSSP5 knockout mouse. The invention additionally encompasses anti-BSSP5 antibodies and methods of production of such antibodies, methods of BSSP5 detection using the antibodies, and the use of BSSP5 proteins or fragments as diagnostic markers for certain medical conditions, e.g., pancreatitis. A method for detecting pancreaticitis comprising measuring BSSP5 concentration in the blood or urine, and a pancreatitis diagnostic agent containing an anti-BSSP5 antibody is also disclosed. Nucleotides encoding BSSP5 were initially isolated in a human brain cDNA library using degenerate PCR primers (AAA61744-A61745) based on conserved regions of serine proteases. The BSSP5 serine proteases and nucleotides encoding them are useful in detecting homologues, mutants and polymorphic variants in biological samples (e.g., blood, urine, brain, prostate gland, placenta, testis, pancreas and spleen) as diagnostic markers for conditions such as Alzheimer's disease, epilepsy, cancer, inflammation, infertility, pancreatitis and prostatic hypertrophy. Sequence AAB11710 represents human BSSP5 (hBSSP5), and sequence AAB11711 represents murine BSSP5 (mBSSP5).

Sequence 264 AA;

Query Match 90.4%; Score 150; DB 21; Length 264;
Best Local Similarity 87.9%; Pred. No. 1.3e-11;
Matches 29; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGGSGGCVPAITPALSYNQR 33
|||||
DB 1 MLLSLTSLVLGGSGGCVPAITPALSFSQR 33

RESULT 3

AAB54077

ID AAB54077 standard; Protein; 192 AA.

AC AAB54077;

DT 09-MAR-2001 (first entry)

DE Human pancreatic cancer antigen protein sequence SEQ ID NO:529.

KW Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
detection; diagnosis; identification; cytostatic; neuroprotective;
nootropic; immunomodulatory; relaxant; contraceptive; gynaecological;
antiinflammatory; cardiant; gene therapy; chromosome mapping;
linkage analysis; tissue identification; tissue typing; forensic;
neural; immune system; muscular; reproductive; gastrointestinal;
pulmonary; cardiovascular; renal; proliferative.

OS Homo sapiens.

PN WO200055320-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US05989.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;

XX WPI: 2000-579444/54.

DR N-PSDB: AAC98842.
XX New nucleic acid that is a pancreatic cancer antigen for preventing,
PT treating, or ameliorating a medical condition, particular pancreatic
PT cancer, or for use in assays for diagnosing a pathological condition -
XX
PS Claim 11; Page 966; 1379pp; English.
XX
CC AAC98773 to AAC99231 encode the human pancreatic cancer associated
CC proteins, called pancreatic cancer antigens, given in AAB54008 to
CC AAB54466. The human pancreatic cancer antigens have cytostatic,
CC neuroprotective, nontoxic, immunomodulatory, relaxant, contraceptive,
CC gynaecological, cardiac and antiinflammatory activities, and can be used
CC in gene therapy. The polynucleotide and proteins can be used for
CC preventing, treating, or ameliorating a medical condition or in assays
CC for diagnosing a pathological condition or a susceptibility to one in a
CC subject. Binding partners to the proteins and the activity of the
CC proteins can be identified. The pancreatic cancer antigens can be used to
CC detect, treat or prevent pancreatic disorders, especially cancer.
CC Agonists and antagonists to the antigens can be screened for. The
CC pancreatic cancer antigen polynucleotides can be used to design nucleic
CC acid hybridisation probes that can be used in chromosome mapping, linkage
CC analysis, tissue identification and/or typing and a variety of forensic
CC and diagnostic methods. The proteins can be used to generate antibodies
CC which are used to purify, detect and target the polypeptides, including
CC both in vivo and in vitro diagnostic and therapeutic methods. The
CC proteins can be used to treat or prevent neural, immune system, muscular,
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 192 AA;
Query Match 80.1%; Score 133; DB 21; Length 192;
Best Local Similarity 86.2%; Pred. No. 1.5e-09;
Matches 25; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 5 SLTSLVLLGSSWGGVPAITPALSYNQR 33
Db 1 SLTSLVLLGSSWGGVPAITPALSYNQR 29
|||||
RESULT 4
AAY64742
ID AAY64742 standard; Protein; 44 AA.
XX
AC AAY64742;
XX
DT 01-FEB-2000 (first entry)
XX
DE Human 5' EST related polypeptide SEQ ID NO:903.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;
KW gene therapy; chromosome mapping; upstream regulatory sequence;
KW forensic; location; development; protein synthesis; stability;
KW regulation; identification.
XX
OS Homo sapiens.
XX
PN WO9953051-A2.
XX
PD 21-OCT-1999.
XX
PF 09-APR-1999; 99WO-IB00712.
XX
PR 09-APR-1998; 98US-0057719.
PR 28-APR-1998; 98US-0069047.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI; 2000-038446/03.
DR

DR N-PSDB: AAZ42356.
XX Novel secreted protein 5' expressed sequence tag sequences used in
PT diagnostic, forensic, gene therapy, and chromosome mapping procedures
PT
PS Claim 3; Page 629; 837pp; English.
XX
CC AAZ42265 to AAZ43075 represent novel 5' expressed sequence tag (EST)
CC sequences, corresponding to human secreted proteins. AAY64651 to
CC AAY65138 represent the EST-related proteins corresponding to AAZ42265 to
CC AAZ43075. The 5' ESTs can be used for producing secreted human gene
CC products. They can be used to identify and isolate 5' untranslated
CC regions (UTRs) and upstream regulatory regions which control the
CC location, development stage, rate, and quantity of protein synthesis, as
CC well as stability of mRNA. The ESTs are also useful as probes for
CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can
CC also be used in forensic procedures to identify individuals, or in
CC diagnostic procedures to identify individuals having genetic diseases
CC resulting from abnormal gene expression. The products may also be used in
CC gene therapy protocols. The nucleic acids encoding signal peptides can be
CC used for directing extracellular secretion of a polypeptide or the
CC insertion of a polypeptide into a membrane, or importing a polypeptide
CC into a cell. The proteins encoded by the EST sequences may be useful in
CC treating a variety of human conditions. Secreted proteins have
CC therapeutic value, and the identification of new secreted proteins is
CC valuable. AAZ42249 to AAZ42264 and AAY64644 to AAY64650 represent
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 44 AA;
Query Match 41.0%; Score 68; DB 21; Length 44;
Best Local Similarity 45.5%; Pred. No. 0.092;
Matches 15; Conservative 4; Mismatches 14; Indels 0; Gaps 0;
QY 1 MLLLSLTSLVLLGSSWGGVPAITPALSYNQR 33
Db 1 MAFLLSCWALLGTFGGVPAIHGCOLSPR 33
|||||
RESULT 5
AAU82738
ID AAU82738 standard; Protein; 263 AA.
XX
AC AAU82738;
XX
DT 23-APR-2002 (first entry)
XX
DE Amino acid sequence of novel human protease #37.
XX
KW Human; protease; cancer; immune-related disorder; cardiovascular disease;
KW neuronal-associated disease; metabolic disorder; inflammatory disorder;
KW nervous system disorder; sexual dysfunction; pain; mood disorder;
KW hypertension; psychotic disorder; neurological disorder; dyskinesia;
KW viral infection; human immunodeficiency virus; HIV; non-viral infection;
KW ocular disease; cytostatic; enzyme.
XX
OS Homo sapiens.
XX
PN WO200200860-A2.
XX
PD 03-JAN-2002.
XX
PF 26-JUN-2001; 2001WO-US20171.
XX
PR 26-JUN-2000; 2000US-214047P.
XX
PA (SUGE-) SUGEN INC.
XX
PI Plowman G, Whyte D, Sudarsanam S, Manning G, Caenepeel S;
PI Charyczak G;
XX
DR WPI; 2002-139913/18.
DR N-PSDB; ABK31780.

XX Nucleic acids encoding novel human proteases, useful for usefull for
 PT treating diseases and disorders such as cancers, immune-related
 PT diseases and disorders, cardiovascular disease (e.g. restenosis) and
 PT inflammatory disorders -
 XX
 XX Claim 6; Fig 2M; 313pp; English.
 XX
 XX The present invention relates to the isolation of novel human
 CC proteases, and the nucleic acids encoding them. The sequences of
 CC the invention are useful for treating diseases and disorders such as
 CC cancers (e.g. breast, colon, lung), immune-related diseases and disorders
 CC (e.g. inflammatory diseases and asthma), cardiovascular diseases
 CC (e.g. restenosis and coronary thrombosis), brain or neuronal-associated
 CC diseases, metabolic disorders (e.g. diabetes, obesity), inflammatory
 CC disorders (e.g. rheumatoid arthritis and psoriasis), central or
 CC peripheral nervous system diseases, migraines, pain, sexual dysfunction,
 CC mood disorders, attention disorders, cognition disorders, hypertension,
 CC hypertension, psychotic disorders, neurological disorders
 CC (e.g. Alzheimer's disease, Parkinson's disease) and dyskinesias.
 CC The nucleic acids and polypeptides are also useful for treating viral
 CC infections caused by human immunodeficiency virus (HIV), and non-viral
 CC infections such as ocular disease (e.g. glaucoma) and macular
 CC degeneration. AAU82702-AAU82760 represent the novel human proteases of
 CC the invention.
 XX
 XX Sequence 263 AA;

Query Match 41.0%; Score 68; DB 23; Length 263;
 Best Local Similarity 51.7%; Pred. No. 0.63;
 Matches 15; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLLGSSWCGVPAITPALS 29
 DB 1 MASLWLLCSFLVGAAGCGVPAIHPVLS 29

RESULT 6
 AAB54191
 ID AAB54191 standard; Protein; 146 AA.
 XX
 AC AAB54191;
 XX
 DT 09-MAR-2001 (first entry)
 XX
 XX Human pancreatic cancer antigen protein sequence SEQ ID NO:643.
 XX
 DE Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
 KW detection; diagnosis; identification; cytostatic; neuroprotective;
 KW nontropic; immunomodulatory; relaxant; contraceptive; gynaecological;
 KW antiinflammatory; cardiant; gene therapy; chromosome mapping;
 KW linkage analysis; tissue identification; tissue typing; forensic;
 KW neural; immune system; muscular; reproductive; gastrointestinal;
 KW pulmonary; cardiovascular; renal; proliferative.
 XX
 OS Homo sapiens.
 XX
 XX WO200055320-A1.
 PN
 XX
 XX 21-SEP-2000.
 XX
 XX 08-MAR-2000; 2000WO-US05989.
 PF
 XX
 XX 12-MAR-1999; 99US-0124270.
 PR
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX
 XX Rosen CA, Ruben SM;
 PI
 XX
 XX WPI; 2000-579444/54.
 DR
 XX
 XX N-PSDB; AAC98956.
 DR
 XX
 XX New nucleic acid that is a pancreatic cancer antigen for preventing,

PT treating, or ameliorating a medical condition, particular pancreatic
 PT cancer, or for use in assays for diagnosing a pathological condition -
 XX
 XX Claim 11; Page 1081; 1379pp; English.
 XX
 XX AAC98773 to AAC99231 encode the human pancreatic cancer associated
 CC proteins, called pancreatic cancer antigens, given in AAB54008 to
 CC AAB54466. The human pancreatic cancer antigens have cytostatic,
 CC neuroprotective, nontropic, immunomodulatory, relaxant, contraceptive,
 CC gynaecological, cardiant and antiinflammatory activities, and can be used
 CC in gene therapy. The polynucleotide and proteins can be used for
 CC preventing, treating, or ameliorating a medical condition or in assays
 CC for diagnosing a pathological condition or a susceptibility to one in a
 CC subject. Binding partners to the proteins and the activity of the
 CC proteins can be identified. The pancreatic cancer antigens can be used to
 CC detect, treat or prevent pancreatic disorders, especially cancer.
 CC Agonists and antagonists to the antigens can be screened for. The
 CC pancreatic cancer antigen polynucleotides can be used to design nucleic
 CC acid hybridisation probes that can be used in chromosome mapping, linkage
 CC analysis, tissue identification and/or typing and a variety of forensic
 CC and diagnostic methods. The proteins can be used to generate antibodies
 CC which are used to purify, detect and target the polypeptides, including
 CC both in vivo and in vitro diagnostic and therapeutic methods. The
 CC proteins can be used to treat or prevent neural, immune system, muscular,
 CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
 CC proliferative disorders: AAC99232 to AAC99240 and AAB54467 represent
 CC sequences used in the exemplification of the present invention.
 XX

SQ Sequence 146 AA;

Query Match 40.4%; Score 67; DB 21; Length 146;
 Best Local Similarity 66.7%; Pred. No. 0.45;
 Matches 12; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 12 LLGSSWCGVPAITPALS 29
 DB 16 LVGAAGCGVPAIHPVLS 33

RESULT 7
 AAY11649
 ID AAY11649 standard; Protein; 195 AA.
 XX
 AC AAY11649;
 XX
 DT 16-JUN-1999 (first entry)
 XX
 XX Human 5' EST secreted protein SEQ ID NO:301.
 DE
 XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
 KW forensic; gene therapy; chromosome mapping; signal peptide;
 KW upstream regulatory sequence; cytokine activity; cell proliferation;
 KW differentiation; haematopoiesis regulation; tissue growth regulation;
 KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; anti-inflammatory; tumour inhibition.
 XX
 OS Homo sapiens.
 XX
 XX WO9906439-A2.
 PN
 XX
 XX 11-FEB-1999.
 PD
 XX
 XX 31-JUL-1998; 98WO-IB01233.
 PF
 XX
 XX 01-AUG-1997; 97US-0904468.
 PR
 XX
 XX (GEST) GENSET.
 PA
 XX
 XX Duclert A, Dumas Milne Edwards J, Lacroix B;
 PI
 XX
 XX WPI; 1999-153700/13.
 DR
 XX
 XX N-PSDB; AAX40367.
 DR
 XX

New polynucleotides which encode polypeptides involved in olfactory sensation for identifying olfactory agonists and antagonists -

Query Match 34.9%; Score 58; DB 22; Length 312;
Best Local Similarity 60.0%; Pred. No. 15;
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 LLLSLTSLVLLGSSWGCGV 21
||||| : | | ||| |||
DB 133 LLLSHRVCLQLAGSACGCV 152

RESULT 12
AAB07181
ID AAB07181 standard; Protein; 441 AA.
XX
AC AAB07181;
DT 22-NOV-2000 (first entry)
XX
DE Drosophila melanogaster imaginal disc growth factor 3.
XX
KW Imaginal disc growth factor 3; IDGF3; chitinase related protein;
KW CHR; tissue growth; wound healing; bone repair; cartilage repair;
KW angiogenesis; meat production; milk production; cancer; gene therapy.
XX
OS Drosophila melanogaster.
XX
PN US6060590-A.
XX
PD 09-MAY-2000.
XX
PF 31-MAR-1998; 98US-0052778.
XX
PR 31-MAR-1998; 98US-0052778.
XX
PA (RECC) UNIV CALIFORNIA.
XX
PI Bryant PJ, Kawamura K;
XX
DR WPI; 2000-349702/30.
XX
CC New imaginal disc growth factor polypeptide, useful for healing wounds,
PT promoting tissue growth, modulating female reproductive tract functions
PT and treating cell proliferative disorders such as cancer and metastasis
PT .
XX
PS Disclosure; Fig 4; 40pp; English.

XX
XX The present sequence is the protein sequence for the Drosophila
CC melanogaster imaginal disc growth factor 3 (IDGF3). The sequence encoding
CC this protein was used to isolate the IDGF4 coding sequence. IDGF4 is a
CC member of the chitinase related proteins (CHRP) and is involved in the
CC promotion of cell growth, motility and morphogenesis. The gene and
CC protein are expected to have mammalian homologues. They can be used in
CC the treatment of cancer, wound healing, tissue regeneration following
CC arthritis, osteoporosis, other skeletal disorders and burns, for
CC revitalising scar tissue resulting from surgical procedures, irradiation,
CC laceration, toxic chemicals, viral or bacterial infection or burns, to
CC promote tissue growth during tissue engineering, for example tissues for
CC skin graft replacements and bone regrowth, and to modulate the function
CC of the female reproductive tract. In addition, they can also be used to
CC increase meat, egg, sperm and milk production in animals. One possible
CC method of use is by gene therapy.
XX
SQ Sequence 441 AA;

Query Match 33.4%; Score 55.5; DB 21; Length 441;
Best Local Similarity 48.3%; Pred. No. 47;
Matches 14; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

QY 2 LLLSLTSLVLLGSSWGCGVPAITPALS 30
||||| : | | ||| |||
DB 210 LLLSLTV-LPNVNSSWYDAPSIAPSLDF 237

RESULT 13
ABB60694
ID ABB60694 standard; Protein; 441 AA.
XX
AC ABB60694;
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 8874.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL04797.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 8874; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 441 AA;

Query Match 33.4%; Score 55.5; DB 22; Length 441;
Best Local Similarity 48.3%; Pred. No. 47;
Matches 14; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

QY 2 LLLSLTSLVLLGSSWGCGVPAITPALS 30
||||| : | | ||| |||
DB 210 LLLSLTV-LPNVNSSWYDAPSIAPSLDF 237

RESULT 14
AAV11650
ID AAV11650 standard; Protein; 123 AA.
XX
AC AAV11650;
DT 16-JUN-1999 (first entry)
XX
DE Human 5' EST secreted protein SEQ ID NO:302.
XX
KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;

KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.

OS Homo sapiens.

XX WO9906439-A2.

XX 11-FEB-1999.

XX 31-JUL-1998; 98WO-IB01233.

XX 01-AUG-1997; 97US-0904468.

XX (GEST) GENSET.

XX Duclert A, Dumas Milne Edwards J, Lacroix B;

XX WPI; 1999-153700/13.

XX N-PSDB; AAX40368.

XX New nucleic acids encoding human secreted proteins - obtained from
XX cDNA libraries derived from liver, lung, large intestine, colon,
XX thyroid and pancreas tissue

XX Claim 27; Page 379-380; 398pp; English.

XX AAX40251 to AAX40397 represent 5' expressed sequence tags (ESTs) for
XX human secreted proteins, and encode the proteins given in AAY11533 to
XX AAY11679, respectively. The proteins given represent the signal peptide
XX and an N-terminal fragment of a secreted protein. The nucleic acid
XX sequences can be used for producing secreted human gene products. They
XX can also be used to develop products for diagnosis and therapy. The
XX proteins obtained may have cytokine activity, cell
XX proliferation/differentiation activity, haematopoiesis regulating
XX activity, tissue growth regulating activity, reproductive hormone
XX regulating activity, chemotactic/ chemokinetic activity, haemostatic and
XX thrombolytic activity, receptor/ ligand activity, anti-inflammatory
XX activity, tumour inhibition activity or other activities. The products
XX can be used in forensic, gene therapy and chromosome mapping procedures.
XX The sequences can also be used for obtaining corresponding promoter
XX sequences. The nucleic acids encoding the signal peptide can be used for
XX directing extracellular secretion of a polypeptide or the insertion of a
XX polypeptide into a membrane, or importing a polypeptide into a cell.

XX Sequence 123 AA;

Query Match 33.1%; Score 55; DB 20; Length 123;

Best Local Similarity 40.7%; Pred. No. 14;

Matches 11; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 3 LLSLTLSLVLLGSGWCGVPATPALS 29

DB 1 MLGITVLAALLACASSCGVFPFNL 27

RESULT 15

AAR67540

ID AAR67540 standard; Protein; 231 AA.

XX AAR67540;

XX 05-AUG-1995 (first entry)

XX Mouse flt-3 ligand.

XX Flt-3 ligand; flt3-L; anemia; cancer; AIDS; gene therapy.

XX Mus sp.

XX Key

XX Location/Qualifiers

XX 1..27

XX Peptide

XX /label= sig_peptide

FT Domain

FT 28..188

FT /label= Extracellular_domain

FT 189..211

FT /label= Transmembrane_domain

FT 212..231

FT /label= Cytoplasmic_domain

XX EP627487-A.

PN 07-DEC-1994.

XX 19-MAY-1994; 94EP-0303575.

XX 24-MAY-1993; 93US-0068394.

PR 12-AUG-1993; 93US-0106463.

PR 25-AUG-1993; 93US-0111758.

PR 03-DEC-1993; 93US-0162407.

PR 07-MAR-1994; 94US-0209502.

PR 11-MAY-1994; 94US-0243545.

XX (IMMV) IMMUNEX CORP.

XX Beckmann MP, Lyman SD;

PI WPI; 1995-008071/02.

XX N-PSDB; AAX79076.

XX Isolated ligands for flt 3 receptors - useful for treating

PT anaemia, AIDS and various cancers

XX Disclosure; Page 25-27; 33pp; English.

XX cDNA encoding mouse flt3-ligand (flt3-L) was isolated from a
XX cDNA library of T-cell line P7B-0.3A4 in CV-1/EBNA-1 cells
XX using a slide autoradiography method. Flt3-L stimulates
XX production of progenitor and stem cells, and can be used e.g.
XX in gene therapy protocols.

XX Sequence 231 AA;

Query Match 33.1%; Score 55; DB 16; Length 231;

Best Local Similarity 61.1%; Pred. No. 27;

Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLLGSSWG 18

DB 192 LLLLLLPLTLVLLAAWG 209

Search completed: December 20, 2002, 15:16:42

Job time : 6.08696 secs